1.3 Zodpovedné používanie UI

Hlavné posolstvo

- Zodpovednosť je úlohou dizajnu, nie kontroly. Detekcia UI nefunguje spoľahlivo.
- A keď detekcia UI zlyhá, je nespravodlivá. Hrozí riziko, že odpovede budú mylne označené ako vygenerované UI, alebo naopak. Viaceré európske krajiny tieto nástroje už vylúčili.
- Sebavedomé vymýšľanie a zdedená zaujatosť vyplývajú z toho, ako sú tieto systémy vytvorené a hodnotené. Dôsledkom by malo byť: učiť o nich, namiesto čakania na opravu.
- Pracovný princíp znie: najprv myslieť, potom UI. Ak ľudia musia najprv sformulovať vlastnú odpoveď, funguje to lepšie než vysvetľovanie fungovania stroja; zabudovanie limitov do softvéru funguje lepšie než zákaz.
- Samotné inštrukcie (prompty) nestačia. Chatbot poverený úlohou tútora má tendenciu skĺznuť späť k poskytovaniu odpovedí, čo robí zo spoľahlivých limitov skôr otázku nákupu (výberu nástroja).
- Zosilnenie (rozdielov) je predvolený stav; vyrovnávanie treba budovať spoločne z prístupu, dizajnu a podpory (scaffoldingu) — a najvzácnejším z týchto troch je vyškolený učiteľ.
- Inklúzia je najjasnejším prínosom a oblasťou, kde učitelia sami vidia najsilnejšie argumenty.

> Ak si z tejto stránky zapamätáte len jednu vetu: Hrádze sa lámu. Učte plávať.

?itate?ský denník, druhý poh?ad

Vráťme sa k nášmu deviatakovi a jeho podozrivo dokonalému čitateľskému denníku. Tento inštinkt je takmer univerzálny a je to inštinkt dobrého učiteľa: zistiť, či použil UI. Postaviť lepšiu hrádzu.

Táto kapitola vysvetľuje, prečo tento inštinkt sám osebe nikam nevedie. Neskôr sa dostaneme k tomu, čo naozaj funguje.

Začnime nepríjemnou časťou. Učitelia nedokážu spoľahlivo rozlíšiť text UI od textu študenta. V nemeckej štúdii dostalo 289 učiteľov s rôznou dĺžkou praxe zmes študentských a chatbotom napísaných textov a mali ich roztriediť. Nedokázala to ani jedna skupina. Napriek tomu si obe skupiny boli isté, že to dokážu. Prax takmer vôbec nezohrala rolu. Vo veľkom meradle je to ešte horšie. Na jednej univerzite výskumníci bežnými kanálmi predložili do reálneho skúškového systému odpovede vygenerované chatbotom, bez toho, aby o tom informovali hodnotiteľov. 94 % z nich nebolo nikdy spochybnených. Navyše boli v priemere hodnotené lepšie než práce skutočných študentov. Aký je teda dôsledok? Siahneme po stroji! A práve tu prestáva byť zistenie len trápne a

stáva sa etickým problémom!

Detektor, ktorý zlyháva u nesprávnych študentov

Výskumníci otestovali sedem najpoužívanejších detektorov UI na esejach TOEFL napísaných študentmi, ktorí sa učia angličtinu ako cudzí jazyk. Popri nich zaradili aj eseje amerických ôsmakov, ktorí sú rodenými hovoriacimi. Práce rodených hovoriacich boli správne klasifikované takmer vždy. Z esejí v cudzom jazyku bolo viac ako šesť z desiatich mylne označených ako napísané strojom.

Mechanizmus nie je žiadnym tajomstvom. Tieto nástroje neodhaľujú prácu veľkých jazykových modelov, ale predvídateľnosť. Text, ktorý používa bežné slová v očakávanom poradí, je vyhodnotený ako umelý. Presne takýto signál produkuje každý, kto píše v druhom jazyku, s menšou slovnou zásobou a bezpečnejšou stavbou viet. Chyby nástroja tak systematicky dopadajú na študentov, ktorí už aj tak nesú najväčšiu záťaž. Inštitúcie si to prepočítali. Univerzita Vanderbilt upozornila, že aj keby sa takýto nástroj mýlil len v 1 % prípadov, pri 75 000 odovzdaných prácach by to znamenalo približne **750 nespravodlivo obvinených študentov ročne**. V dôsledku toho ho vypli. Usmernenie nemeckého spolkového ministerstva školstva je jednoznačné: technológia *„nie je dostatočne spoľahlivá na to, aby poskytla právne bezpečný dôkaz.“*

Veľká časť Európy o tom už prestala diskutovať. Francúzsko, Holandsko a Švajčiarsko odrádzajú od používania detektorov alebo ich zakazujú. Britský skúškový regulátor sa zámerne rozhodol uprednostniť ľudský úsudok pred softvérom. Holandský argument ide o krok ďalej a stojí za zapamätanie: vloženie práce študenta do detektora je samo osebe spracovaním jeho osobných údajov, a robiť to bez súhlasu porušuje zákon o ochrane osobných údajov. → 1.4 Súlad s predpismi pri používaní UI

Zdá sa, že panika okolo podvádzania je slabšia než hluk okolo nej. Výskumníci sa pýtali študentov na troch stredných školách pred vznikom ChatGPT aj po ňom. Miera podvádzania zostala vo veľkej miere rovnaká. Väčšina študentov odmietala nechať chatbota, aby za nich jednoducho vytvoril prácu, no považovali za spravodlivé použiť ho na začatie práce alebo na vysvetlenie niečoho.

Skutočná otázka nikdy neznela: „Ako ho chytím?“ Znie: čo ešte domáca úloha meria, keď jej výsledok je možné vygenerovať na požiadanie?

Dve veci, ktoré sa nedajú opravi?, a preto sa musia u?i?

Sebavedomé vymýšľanie je zabudovanou vlastnosťou. Tieto systémy generujú pravdepodobné pokračovanie textu, čo nie je to isté ako pravdivé pokračovanie. Výskumníci teraz formálne preukázali, odkiaľ to pochádza: štandardné testy používané na hodnotenie veľkých jazykových modelov odmeňujú sebavedomé hádanie viac než priznanie neistoty, a pokiaľ sa modely takto

hodnotia, pretrváva určitá miera plynulých, dobre sformulovaných nepravdivých tvrdení. Je to vlastnosť spôsobu, akým sú tieto systémy vytvorené a hodnotené, nie chyba čakajúca na opravu.

Zabudované sú aj skreslenia z trénovacích dát. Navyše sa nezmenšujú s tým, ako sa technológia zlepšuje. Európska štúdia rodových stereotypov naprieč jazykmi zistila, že väčšie modely, aj po dodatočnom trénovaní zameranom na to, aby boli bezpečnejšie a spravodlivejšie, niekedy stereotypizujú viac, nie menej. Ak sa vlastnosť nedá odstrániť technicky, stáva sa súčasťou učiva. A dá sa naučiť: v jednom fínskom programe sa podiel študentov, ktorí dokázali rozpoznať zaujatosť vo výstupe systému UI, zvýšil zo 7 % na 44 %. Dôsledkom pre vyučovanie musí byť trénovanie analýzy chýb namiesto ich vyhýbania sa. Inými slovami: chyby stroja sa stávajú súčasťou lekcie.

?o funguje: dizajn, nie kontrola

Dôkazy poukazujú na jeden jednoduchý a ľahko zapamätateľný princíp: najprv myslieť, potom UI.

- Najprv si vyžiadať záväzok od človeka. V sérii experimentov ľudia, ktorí si museli najprv zapísať vlastnú odpoveď skôr, než videli návrh stroja, odhalili oveľa viac jeho chýb než ľudia, ktorým boli ukázané podrobné vysvetlenia toho, ako stroj uvažoval. Vysvetľovanie UI problém nevyriešilo, no vynútenie vlastného úsudku áno. S jedným úprimným upozornením: verzie, ktoré fungovali najlepšie, sa účastníkom páčili najmenej, a najviac pomohli tým, ktorí už aj tak radi premýšľajú s väčším úsilím. Ide teda o jav známy aj ako Matúšov efekt: tomu, kto má, sa ešte pridá. Práve títo z tohto prístupu profitujú najviac.
- Zabudovať limity priamo do nástroja. Chatbot z Kapitoly 1.2, ktorý namiesto odpovedí kládol otázky, je tá istá myšlienka premietnutá do softvéru: študenti, ktorí ho používali, na skúške nič nestratili, zatiaľ čo tí s neobmedzenou verziou dopadli zle.
- Robiť proces viditeľným namiesto naháňania výsledku. Návrhy, pracovné poznámky, krátky rozhovor o práci. Tie ukážu, čo študent dokáže, bez toho, aby niekto musel dokazovať, čo urobil.
- Dánsko tento princíp pilotne testuje: na zúčastnených školách môžu študenti používať UI pri príprave na ústnu skúšku z angličtiny, zatiaľ čo písomná skúška obsahuje ručne písanú časť bez pomôcok.

Predtým, než sa na to spoľahnete, stojí za to poznať dve obmedzenia. Prvé: populárna rada, že učitelia majú chatbotu jednoducho nariadiť, aby sa správal ako sokratovský tútor, neobstojí pri dlhšej konverzácii. Aspoň v súčasnosti má systém tendenciu skĺznuť späť k poskytovaniu odpovedí. Spoľahlivé limity musia byť zabudované priamo do systému, čo z nich robí skôr otázku nákupu než zručnosti pri formulovaní promptov. Druhé: aby UI vysvetlila svoje uvažovanie, znie ako samozrejmá poistka, no vysvetlenia spoľahlivo nezabraňujú nadmernej dôvere. V jednej štúdii nezmenili mieru, akou ľudia prijímali nesprávne rady, a vysvetlenia, ktoré sú príliš technické alebo príliš zjednodušené, môžu nadmernú dôveru dokonca zvýšiť. Vysvetlenie je podmienkou, nie riešením.

Postoj: dva režimy zlyhania, nie jeden

Obávame sa, že učitelia dôverujú UI príliš. Rovnako by nás však mal znepokojovať aj opak. Aj skúsení ľudia môžu byť chytení do pasce nadmernej dôvery. V jednej štúdii učitelia kontrolovali známky vygenerované UI. 42 % spätnej väzby stroja hodnotili ako nejasnú alebo nesprávnu, no zmenili len 9 % z nej. Odhaliť chybu a opraviť ju sú dva odlišné procesy.

Opačným pólom je zatajovanie. Učitelia skrývajú svoje vlastné používanie UI pred kolegami a študentmi a zakazujú im to, čo sami robia: 77 % z nich používa UI súkromne, zatiaľ čo 57 % tvrdí, že študenti by ho nikdy nemali používať na generovanie nápadov. Zároveň štúdia akademicky nadaných študentov v 6. až 8. ročníku zistila, že medzi tretinou až dvoma pätinami z nich už tvrdí, že UI vie viac než ich učitelia. Profesia, ktorá skrýva svoje postupy, nemôže byť vzorom zodpovednej praxe, a transparentnosť nie je len zdvorilosťou. Je to mechanizmus, vďaka ktorému funguje učenie sa príkladom.

Vyrovnáva UI podmienky, alebo ich naklápa?

Pri pracovných úlohách UI vyrovnáva rozdiely. V jednom experimente dosahovali účastníci s vyšším vzdelaním výrazne lepšie výsledky než tí s nižším vzdelaním. Keď každý dostal asistenta UI, tieto rozdiely takmer úplne zmizli. Keď sa asistent opäť odobral, veľká časť rozdielov sa vrátila. To, čo UI uzavrela, bola medzera vo výstupe, nie medzera v schopnostiach.

Pri učebných úlohách je to presne naopak. Jedna štúdia nezistila v priemere žiadny prínos, ale narastajúci odstup medzi študentmi, ktorí už vedeli veľa, a tými, ktorí nevedeli. Iná štúdia sa zamerala na študentov píšucich v cudzom jazyku: slabší autori dostávali od UI viac opráv, no úspešne využili menej z nich, a mnohí považovali záplavu spätnej väzby za skľučujúcu. Spätná väzba bez schopnosti ju roztriediť nepomáha, ale zahlcuje.

Táto medzera existuje ešte predtým, než sa niekto vôbec dotkne nástroja. V Nemecku vníma UI ako príležitosť 80 % študentov z najviac zvýhodnených rodín, oproti 55 % z najmenej zvýhodnených. Vo Veľkej Británii nevedie deliaca čiara cez sklad zariadení, ale cez zborovňu: 45 % učiteľov na súkromných školách absolvovalo formálne školenie v používaní UI, oproti 21 % na štátnych školách. Aj v rámci štátneho sektora bohatšie školy opäť predbiehajú tie chudobnejšie. Prieskum medzi deťmi v 20 európskych krajinách zisťuje rovnaké sociálne rozdelenie v tom, ako často a akým spôsobom UI používajú. Jedna veta vystihuje celý problém lepšie než akékoľvek číslo: *„bohatí majú prístup k technológii aj k ľuďom, ktorí im pomôžu ju používať, zatiaľ čo chudobní majú prístup len k technológii.“* Je to najsilnejší dostupný argument v prospech toho, že učitelia sú kľúčoví na zmenšenie tejto medzery.

Kľúčovým obmedzujúcim faktorom teda je, či má žiak učiteľa, ktorý bol na to vyškolený. UI kompenzuje rozdiely vtedy, keď spolu fungujú tri veci: bezpečný prístup, nástroj navrhnutý na učenie, nie na poskytovanie odpovedí, a učiteľ, ktorý jeho používanie sprevádza. Ak čo i len jedno z toho chýba, UI namiesto toho rozdiely zosilňuje. Keďže všetky tri prvky sa zvyčajne vyskytujú spoločne na tých istých školách — a rovnako spoločne chýbajú na tých istých školách — zosilňovanie je predvolený stav a vyrovnávanie je niečo, čo treba aktívne vybudovať.

Revision #14
Created 2026-09-01 09:13:36 UTC by Alexander
Updated 2026-09-03 08:17:48 UTC by Alexander