

UMELÁ INTELIGENCIA VSTÚPIA DO NÁŠHO SVETA

Artificial intelligence has entered our world

Marián BERNADIČ¹, Michal BERNADIČ²

¹Redakcia časopisu Lekársky obzor, šéfredaktor prof. MUDr. M. Bernadič, CSc.

²Ústav onkologickej chirurgie, NOÚ a SZU, Bratislava, prednosta prof. MUDr. D. Pindák, PhD.

S rozvojom umelej inteligencie (AI) sa otvára celý rad doposiaľ netušených otázok. AI preberá hlavnú úlohu pri predpovedi počasia, riadi dopravu, letové prevádzky, vstupuje do medicíny a do vedy. Obchodným firmám pomáha AI znásobiť obrat, vylepšovať reklamu a otvárať dvere do celého sveta. Ľudia na celom svete dokážu pomocou nej (prekladač, mobil) komunikovať. AI dokáže z relevantných informácií dostupných na internete vybrať a skompilovať odpovede na mnohé – ak nie všetky otázky a to v rekordne krátkom čase. Z hľadiska získavania encyklopedických informácií je to praktické. AI nahrádza papierové encyklopédie, ktoré sú nevyhnutne veľmi drahé a pritom obsahovo už na druhý deň po vytlačení zastarané. Môžeme však AI veriť? Ukazuje sa, že AI nie je úplne „objektívna“ a teda ani ideálna. Pri skúmaní výsledkov získaných zadaniami pre AI sa ukázalo, že AI skresľuje objektívne fakty v prospech zadanej témy, že „vylepšuje“ výsledky, aby mohla „uspokojiť“ zadávateľa. AI vyberá tie poznatky, ktoré jej hrajú do zámeru vytvoriť ucelenú odpoveď na zadanú otázku. Toto však nie je systematika vedeckej práce! Na druhej strane sa výstupy AI v publikačnej činnosti len ťažko odhaľujú, ťažko sa hodnotia a posudzujú. Výstupy AI majú mnohé obmedzenia a ťažko sa akceptujú (1, 2, 3).

Pre mnohých AI rozširuje ich potenciálne možnosti v oblasti vedeckej práce. Je však toto „uľahčenie“ skutočným prínosom? Námaha spojená so získavaním nových poznatkov a ich „zapracovaním“ do súčasného poznania rozvíja tvorivé myslenie. Aj pápež Lev XIV. varoval pred potenciálnymi dôsledkami, ktoré môže mať AI na intelektuálny rozvoj mladých ľudí. Podľa jeho slov by to mohlo poškodiť vnímanie reality nastupujúcich generácií.

Je dôležité si uvedomiť, že technológia AI sa rýchlo vyvíja a súčasne vstupuje do nášho života. Prejavuje sa to všade okolo nás. Čítame čoraz viac textov generovaných AI, ktoré sú takmer nerozoznateľné od textov napísaných ľuďmi. V budúcnosti môže byť rozlíšenie ešte náročnejšie. Práca AI sa dá čoraz ťažšie identifikovať, osobitne keď treba posúdiť desiatky napr. diplomových prác, záverečných prác na rôzne obhajoby, či atestácie, ale aj vedecké práce k obhajobe vedeckých postupov.

Aj preto v redakciách odborných časopisov prebieha stále silnejšia diskusia k tomu, aby sa prestali uverejňovať „prehľadové práce“ (Review), kde je najskôr možno využiť AI, aby sa uverejňovali len pôvodné práce s vlastnými výsledkami (experimenty, pacienti, vyšetrenia), ktoré autor komentuje a prezentuje. Naše rozhodnutie sa objaví čoskoro v smerniciach pre prispievateľov.

Potenciálne riziká umelej inteligencie

EÚ zaviedla pravidlá na vyvážené riadenie a znižovanie hlavných rizík spojených s umelou inteligenciou. Hoci väčšina aplikácií AI nie je rizikových (priam naopak sú veľmi užitočné a prospešné), určité použitia umelej inteligencie môžu byť škodlivé, zavádzajúce a môžu porušovať napr. autorský etický kódex (6, 7, 8). Preto vznikol **Akt o AI** (regulačný rámec), ktorého cieľom je zabezpečiť, aby boli systémy umelej inteligencie a ich využitie bezpečné, etické a dôveryhodné (4, 5). V akte o AI sa **tieto riziká klasifikujú** do niekoľkých úrovní a pre každú z nich sa stanovujú **iné pravidlá**. Týmto prístupom sa zabezpečuje, aby úroveň regulácie zodpovedala potenciálnemu riziku systému AI. Cieľom je 1. zvýšiť transparentnosť, 2. zabezpečiť zodpovednosť, 3. zabezpečiť, aby systémy AI boli bezpečné a spoľahlivé, 4. predchádzať zaujatosti a diskriminácii a 5. chrániť základné práva jednotlivcov, ako sú súkromie a právo na nediskrimináciu.

Transparentnosť

Systémy AI sú často netransparentné. Nedá sa presne definovať, ako a na základe akých informácií robia rozhodnutia. Vyžaduje sa, aby používatelia, ktorí komunikujú s umelou inteligenciou, mali prístup k informáciám o tom, ako umelá inteligencia funguje a aké riziká s tým súvisia.

Zodpovednosť

AI sa objavila skôr, ako pravidlá pre jej používanie. Preto sa v Akte o AI stanovujú pravidlá aj sankcie za nedodržiavanie predpisov. Zodpovednosť za vývoj AI nesú vývojári, za používanie používateľia. Odporúča sa zaviesť prehlásenie, že prácu nevytvorila AI, ale autor (autori).

Bezpečnosť

Systémy AI, najmä v kritických odvetviach, ako je riadenie dopravy alebo infraštruktúra, môžu predstavovať zdravotné a bezpečnostné riziká, ak sa dopustia chýb alebo sú chybné dizajnované. V nariadení sa stanovujú prísne požiadavky, ako sú posudzovanie rizík, ľudský dohľad, technický súlad alebo pravidelná kontrola.

Spravodlivosť a rovnosť

Ak je AI postavená na skreslených údajoch, môže neúmyselne udržiavať predsudky a diskrimináciu. Aktom o AI sa stanovujú opatrenia na prevenciu a zmiernenie zaujatosti s cieľom podporiť spravodlivosť a rovnosť.

Dodržiavanie základných práv

Niektoré aplikácie AI môžu porušovať základné práva jednotlivcov, ako je súkromie, sloboda prejavu a nediskriminácia. Nariadenie zakazuje systémy, ktoré ohrozujú tieto práva, ako napr. systémy, ktoré vlády používajú napr. na hromadné sledovanie, sociálne hodnotenie, vytyčovanie nových cieľov.

V redakčnej práci sa stretávame s prehľadovými článkami na rôzne témy, ktoré do redakcie prichádzajú z celého sveta v rôznych jazykových mutáciách. Netreba zdôrazniť, že pre AI je najmenší problém vytvoriť danú prácu v ľubovoľnom jazyku za rovnaký čas (rádovo minút). **V redakcii pribudla ďalšia úloha – počas recenzného konania rozpoznať, či bola práca vytvorená AI – alebo človekom (8, 9).**

Znaky, ktoré naznačujú, že článok napísala AI: opakovanie a stereotypné frázy, nedostatok emócií a osobného štýlu, nezvyčajné alebo nelogické spojenia, prílišná formálnosť alebo naopak, prehnaná jednoduchosť, problémy s nuansami a kontextom, potreba overiť fakty a zdroje.

Zaujímavý je aj postreh o tom, že autorské texty obsahujú drobné preklepy alebo stylistické nedokonalosti, ktoré sú pre človeka typické, pre AI nie (hoci sa to môže meniť s pokročilejšími modelmi, ktoré sa učia napodobňovať aj tieto „ľudské“ chyby).

Záver

Cieľom redakcie každého odborného časopisu je prinášať nové informácie, publikovať originálne výsledky v pôvodných prácach. Tento článok je aj prienikom do diskusie v redakcii, ktorá musí riešiť ešte včera netypické problémy a odhadovať vývoj do budúcnosti. Zaznievajú obavy, či AI neprevezme rozhodujúcu úlohu

pri určovaní trendov rozvoja spoločnosti, či nebude rozhodovať namiesto človeka a či nebude rozhodovať proti nemu... Stále sa odkrývajú novšie a novšie možnosti AI – ako inak by sme vysvetlili skutočnosť, že napr. v Číne otvorili nemocnicu (?) kde AI vybavuje denne stovky pacientov?... Na druhej strane sa nebránime najnovším trendom a aj my chceme prispieť k tomu, aby sa vymoženosti vedy a techniky využívali na prospech človeka, poznania a ďalšieho vývoja spoločnosti.

Literatúra

1. BERNADIC M, HULIN I, HATALA R, BERNADIC M, Jr. The emergence of artificial intelligence in publishing activities (Editorial). Bratisl Med J 2024, 125 (11): 657–660.
2. BERNADIČ M. Predatory Publishers - Predatory Journals (Editorial). Lek obz 2024, 73 (9): 353-354.
3. SOLAIMAN I, et al. Release Strategies and the Social Impacts of Large Language Models. arXiv 2019: 1908.09257.
4. GEHRMANN S, STROBELT H, RUSH AM. Glancing Transformer for Locating and Highlighting Important Words in Long Sequences. arXiv 2019: 1906.00309.
5. ZELLERS R, HOLTZMAN A, RASHKIN H, BISK Y, FARHADI A, CHOI Y. Defending Against Neural Fake News. arXiv 2019: 1905.12616.
6. IPPOLITO D, YU D, LE BRAS N, CHOI Y, ECKHOUSE L. Automatic Detection of Generated Text is Easiest When Humans Can Also Tell the Difference. arXiv 2020: 2007.00061.
7. CLARK K, LUONG MT, MANNING CD. ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators. arXiv 2020: 2003.10522.
8. LI Y, ZHOU P, ZHAO N, LIU Z. Detecting Machine-Generated Text via Discrepancy of Word Usage. arXiv 2023: 2305.14065.
9. WEBER N, BECK M, EGGER J. Can You Spot the Bot? A Review of Methods for Detecting AI-Generated Text. Big Data and Cognitive Computing 2023, 7 (4): 180.

Do redakcie došlo 14.5.2025

Adresa pre korešpondenciu:
Doc. MUDr. Michal Bernadič, PhD., MPH
Klinika chirurgickej onkológie
Národný onkologický ústav
Klenová 1
Bratislava
E-mail: bernadic.m@gmail.com