

V dlhodobých cieľoch ako zvyšovanie dĺžky života, zlepšovanie kvality ovzdušia alebo aj v oblasti rodovej rovnosti sa často umiestňujeme na spodných pozíciách štatistík EÚ. Napriek tomu sa nám podarilo zviditeľniť jednou svetovou novinkou v oblasti umelej inteligencie použitou pri volebnej kampani.¹ Nie je to však až taká pozitívna správa – išlo o diskreditačný deepfake. Pri očakávaní volieb v ďalších krajinách sa aj vďaka našej skúsenosti čoraz častejšie vynára jedna otázka v rôznych obmenách: Sme pripravení na umelú inteligenciu?

S rozširovaním využívania umelej inteligencie sme nútení čeliť novým hrozbám v podobe dezinformácií. Bojujú proti nim viaceré organizácie po celom svete, na Slovensku k nim patrí projekt Demagóg, ktorý overuje správnosť tvrdení politikov. Stojí pred náročnou výzvou, ktorou je overovanie dezinformácií v novej podobe. Pritom sa rýchlo stávajú jednými z najnebezpečnejších a vyvrátiť ich v reálnom čase je takmer nemožné. Znepokojenie nielen na Slovensku vyvolala forma predvolebnej falošnej audionahrávky vytvorenej umelou inteligenciou, keďže jej overenie trvalo dlhšie a bolo náročnejšie ako overenie videa. Zároveň správy v tejto forme pôsobia oveľa presvedčivejšie ako len písaný text, takže majú väčšiu šancu stať sa internetovou senzáciou. Aj keď máme zaručenú slobodu prejavu, neznamená to, že máme právo vedome vytvárať a šíriť dezinformácie. Síce niektorí experti² zatiaľ nepredpokladajú, že v najbližších slovenských voľbách sa budú masy voličov rozhodovať na základe falošných informácií z deepfake videí, situácia o niekoľko rokov môže byť úplne iná. Pri voličoch, ktorí nie sú pevne rozhodnutí, koho budú voliť, je už teraz vyššia šanca, že prehodnotia svoj názor na základe dezinformácie. Napriek vyjadreniam „zúčastnených“ osôb, že ide o falošné dôkazy proti nim, môže video či audionahrávka spôsobiť veľké škody na ich reputácii. Preto je dôležitý rýchly zásah zo strany sociálnych sietí už v prvých hodinách po zverejnení a stiahnutie materiálu z internetu, čím sa podniknú aspoň v istej miere kroky na ochranu práva obete na zachovanie dobrej povesti³. Nanešťastie internet nezabúda. Vyvrátiť „presvedčivé“ dôkazy a úplne zabrániť ich šíreniu na rôznych platformách je v praxi nemožné. Navyše so zlepšujúcimi sa technológiami je stále náročnejšie rozlíšiť realitu od fikcie.

Na vlastnej koži sme zažili jedno z prvých, no určite nie posledných zneužití umelej inteligencie na politické účely. Týmto incidentom sa poukázalo na dôležitú a obsérnu tému, o ktorej treba diskutovať vo svete a aj u nás. Umelá inteligencia nám do životov zasahuje už dlhšiu dobu, no až v takýchto vyhrotených situáciách si to naozaj začíname uvedomovať aj my, nie až tak zorientovaní laici. Kde sú hranice, v ktorých je etické využívať umelú inteligenciu? A kedy sú tieto pomyselné hranice prekračované a zasahujú do našich práv? Otázka etickosti je v niektorých prípadoch jednoduchšia, napríklad už spomínaná tvorba a šírenie dezinformácií jednoznačne zasahujúcich do práva na informácie.⁴ Európska únia prijíma ako prvá na svete komplexné legislatívne kroky na kontrolu umelej inteligencie,⁵ vďaka ktorým by mali byť naše základné práva chránené. Rôzne modely umelej inteligencie rozdelila do skupín a vyhodnotila úroveň rizika z hľadiska ľudských práv. Používanie niektorých z nich nebude v budúcnosti v Európskej únii povolené.

Zakázané budú modely priamo zasahujúce do súkromia osôb, ako napríklad systém biometrického rozpoznávania tváří či rozpoznávania emócií. Poznať presné ľudské emócie v konkrétnych situáciách by znamenalo poznať aj ich politické či náboženské presvedčenia a aj sexuálnu orientáciu. Informácie tohto typu by nemali byť prístupné nadriadeným ani iným osobám bez vedomia a súhlasu jednotlivca. Tieto dáta nepatria do pracovnej či verejnej sféry

a narábanie s nimi je vážnym zásahom do súkromia. Musíme však robiť viac aj pre ochranu súkromia pred stratou anonymity používateľských účtov. Vďaka umelej inteligencii je identifikácia osôb na internete jednoduchšia a informácie, ktoré sme považovali za zabezpečené, môže stránka prepojiť s ďalšími údajmi a tie následne predávať tretej strane.

Aj k všeobecnej regulácii prichádzajúcej z EÚ⁵ však musíme podľa môjho názoru pridať opatrenia vo viacerých oblastiach, akou je napríklad vzdelávanie. So sprístupnením umelej inteligencie verejnosti sme vstúpili na neprebádané územie, ktoré najrýchlejšie objavujú študenti. Často nevidia dôvod na zdĺhavé hľadanie podkladov pre svoju esej či záverečnú prácu, keď im stačí opýtať sa textového modelu umelej inteligencie – chatbota. Má zmysel prehl'adávať rôzne stránky a čítať desiatky prác a článkov, keď si vieme vypýtať zhrnutie bez väčšej námahy a dostaneme ho v pomerne dobrej kvalite? Ak by sme chceli zájsť ešte ďalej, pýtali by sme sa samých seba otázku, či má vôbec zmysel písať takúto prácu, ak ju za nás na vysokej úrovni vie napísať niekto iný. Spoznať rozdiel medzi odborným článkom napísaným človekom a textovým modelom umelej inteligencie dnes nie je také ľahké, ako by sme si radi mysleli. Aj z tohto dôvodu sa niektoré školy rozhodli učiť žiakov, či už na hodinách alebo na krúžkoch, ako využívať túto modernú technológiu pre svoj prospech, no zároveň s rozumom. Vďaka proaktívnym učiteľom budú študenti vedieť, že informácie od chatbotu si treba ďalej overiť a nielen im slepo dôverovať. Ale čo školy, ktoré k takémuto prístupu študentov nevedú, pretože oficiálne nemusia? Nebudú ich absolventi na pracovnom trhu znevýhodnení oproti tým, ktorí správnym použitím umelej inteligencie dosahujú výsledky rýchlejšie a presnejšie? Ústava nám garantuje právo na **kvalitné** vzdelanie⁶, a aby sme sa k takému dopracovali, musíme v školstve aktívne reagovať na zmeny vo svete technológií aj na trhu práce, pretože ak tak neurobíme, budeme pripravovať obrovské množstvá ľudí na pozície, o ktoré bude klesať záujem zo strany zamestnávateľov.

Ďalšou oblasťou, v ktorej našla umelá inteligencia uplatnenie už pred nejakým časom sú odporúčacie algoritmy sociálnych sietí, ale aj rôznych stránok, ktoré nám podľa toho zobrazujú buď články, alebo reklamy, o ktorých predpokladajú, že by nás mohli zaujať. Robia tak na základe metadát, ktoré im o sebe väčšinou nevedome poskytujeme, a na základe našich vyhľadávaní. Následne sa používajú rôzne metódy ako napríklad „najbližší sused“,⁷ ktoré majú za úlohu nás zaujať zobrazovaním obsahu, na ktorý reagovali podobní užívatelia ako my. Utvrdzujeme sa tak v našom svetonázore v niektorých prípadoch až do takej miery, že strácame schopnosť objektívne vyhodnocovať informácie a argumenty podporujúce iné názory. Stráca sa dialóg medzi ľuďmi s rozdielnymi pohľadmi na svet, a tak sa nemôžeme čudovať novodobému fenoménu, akým je vysoká miera polarizácie v našej spoločnosti.

Medzi výzvy, ktorým čelíme pri algoritmoch patrí aj „algoritmická diskriminácia“ spôsobená algoritmickým skreslením.⁸ Umelá inteligencia nie je síce spočiatku ovplyvnená rovnakými predsudkami, aké má človek, ale vytvára si svoje vlastné. Jedným zo základných systémov jej učenia je strojové učenie, pri ktorom v tréningovej množine sú prechádzané veľké množstvá dát a vytvárajú sa vlastné spojenia medzi pojmami a charakteristikami. Kameňom úrazu sú dáta, ktoré pri tréningu konzumuje. Pokiaľ v nej vedci úmyselne alebo omylom umiestnia jednostranne zamerané informácie, umelá inteligencia si na ich základe vytvorí model správania. Ak by sme ju chceli použiť na rozhodnutie o vine alebo nevine a pri tréningu sa stretávala prevažne s rozhodnutiami o vine v súvislosti s istou etnickou alebo inou skupinou, nie je vylúčené, že to ovplyvní jej rozhodnutie v neprospech obžalovaného z takejto skupiny.

Ako dôkaz obmedzeného fungovania algoritmov umelej inteligencie môžeme použiť štúdiu,⁹ v ktorej sa vedci umelej inteligencie pýtali, ako by vyriešila diplomatický konflikt. Všetky modely riešili situáciu viac či menej agresívnym spôsobom. Hypotéza, ktorou sa toto správanie vysvetľuje, hovorí o prevažujúcom množstve prác analyzujúcich ozbrojené konflikty, a nie mierové riešenia – konanie umelej inteligencie bolo založené aj na dátach v istom zmysle propagujúcich násilie. Ak teda nájde na základe dát prepojenia medzi istou skupinou a istým správaním, môže dôjsť k diskriminácii v jej rozhodovaní bez toho, aby sme si to všimli. Tieto algoritmy nie sú dostatočne transparentné pre ich označovanie ako obchodné tajomstvo. Verejnosť tak nemá prístup k presným odôvodneniam rozhodnutí umelej inteligencie, ale je pravda, že niekedy si nemôžu byť stopercentne istí ani vedci, na základe ktorých prepojení sa v konkrétnych situáciách rozhoduje.

Zatiaľ som písala o ľudských právach, ale mám pocit, že nespomenúť aj práva umelej inteligencie by nebolo spravodlivé, pretože v priebehu najbližších rokov musí byť vo verejnom priestore diskusia aj na túto tému. Na internete vidíme množstvo obsahu vytvoreného s jej pomocou. Podľa nových regulácií EÚ bude povinné uviesť, kto je autorom diela – či človek, alebo umelá inteligencia.⁵ Môžeme teda povedať, že ide o prvé čiastočné priznanie jej práv, aspoň tých autorských. O úplnom priznaní nemôžeme hovoriť, pretože nie je jasné, kto by z jej diela mal profitovať. Mohla by to byť osoba, ktorá jej zadala správny príkaz, alebo firma, ktorá model sprístupnila. Aby sa mohlo duševné vlastníctvo skutočne celé pripísať umelej inteligencii, musela by dosiahnuť istý stupeň samostatnosti v rozhodovaní a začať si uvedomovať svoju existenciu. Zatiaľ sa nám nepodarilo vytvoriť všeobecnú umelú inteligenciu, ktorá by vedela konať autonómne, a s overením, či vie o svojej existencii to tiež nie je také jednoduché. Stále totiž nevieme ani to, ako presne uvedomovanie si seba samého, teda vedomie funguje u ľudí. Kým sa nám nepodari objasniť aspoň základy fungovania toho ľudského, diskusia o vedomí umelej inteligencie bude skôr hádaním ako naozajstnou vedeckou diskusiou na vysokej úrovni. S pokrokom vo výskume oboch oblastí – psychológie aj umelej inteligencie – však verím, že sa nám podarí nájsť ďalšie riešenia autorských práv jednoduchších modelov a diskutovať o širšom rozsahu potenciálnych práv pre všeobecnú umelú inteligenciu.

Napriek možným rizikám a výzvam spojených s umelou inteligenciou by nebolo správne prestať ju vyvíjať a využívať. Jej zaradenie do bežného života je významným pokrokom v porovnaní s minulosťou a má potenciál uľahčiť nám život. Je na nás, aký prístup zvolíme. Môžeme sa tváriť, že ďalšia technologická revolúcia nenastala, a pred touto problematikou zatvárať oči. To by ale znamenalo nezanedbateľné zaostávanie v porovnaní s ostatnými krajinami, ktoré sa ju rozhodnú využívať. Môžeme sa vrhnúť do jej rozvoja a čo najrýchlejšie ju zapojiť do každej možnej oblasti s vierou, že všetky relevantné riziká budú čoskoro odstránené. Museli by sme si určiť, s akou presnosťou jej rozhodnutí sa uspokojíme. Ja osobne momentálne vidím priveľa nedostatkov na zverenie väčšej zodpovednosti umelej inteligencii, a preto by som zvolila prístup niekde medzi oboma extrémami a minimálne najbližších pár rokov presadzovala jej implementáciu len do niektorých oblastí, prípadne len na niektoré vybrané pozície, kým sa systémy nestanú spoľahlivejšími. Možno zvolíme ako spoločnosť iný prístup, ktorý bude výhodnejší, a dnes si ho ešte neviem ani ja, ani väčšina ľudí predstaviť ako možnosť. Nech už sa rozhodneme akokoľvek, nesmieme zabúdať na to, že umelá inteligencia je zatiaľ len nástroj, a či ho využijeme dobre, záleží na nás, na ľudskej inteligencii.

Zdroje a použité študijné materiály:

¹ MEAKER, M. *Slovakia's Election Deepfakes Show AI Is a Danger to Democracy*. In *Wired*. [cit. 22.2.2024] Dostupné na internete: < <https://www.wired.co.uk/article/slovakia-election-deepfakes> >.

² HLAVANDA GOGOVÁ, R. *Deepfake ráta s tým, že naša pozornosť je krátka. V kampani to môže ovplyvniť nerozhodnutých, tvrdí expert*. In *Hospodárske noviny*. [cit. 22.2.2024] Dostupné na internete: < <https://hnonline.sk/tv/relacie/doslova/96132137-deepfake-rata-s-tym-ze-nasa-pozornost-je-kratka-v-kampani-to-moze-ovplyvniť-nerozhodnutych-tvrdi-expert> >.

³ 460/1992 Zb. Ústava Slovenskej republiky, Druhá hlava, Druhý oddiel, Čl. 19 [cit. 22.2.2024] Dostupné na internete: < [460/1992 Zb. - Ústava Slovenskej republiky - SLOV-LEX](#) >.

⁴ 460/1992 Zb. Ústava Slovenskej republiky, Druhá hlava, Tretí oddiel, Čl. 26 [cit. 22.2.2024] Dostupné na internete: < [460/1992 Zb. - Ústava Slovenskej republiky - SLOV-LEX](#) >.

⁵ *Umelá inteligencia*. In *Európska rada a Rada EÚ*. [cit. 22.2.2024] Dostupné na internete: < <https://www.consilium.europa.eu/sk/policies/artificial-intelligence/> >.

⁵ Rada EÚ. "Akt o umelej inteligencii: Rada a Parlament dosiahli dohodu o prvých pravidlách na svete pre umelú inteligenciu" Tlačová správa rady EÚ, 9.12.2023. In *Európska rada a Rada EÚ*. [cit. 22.2.2024] Dostupné na internete: < <https://www.consilium.europa.eu/sk/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> >.

⁶ 460/1992 Zb. Ústava Slovenskej republiky, Druhá hlava, Piaty oddiel, Čl. 42 [cit. 22.2.2024] Dostupné na internete: < [460/1992 Zb. - Ústava Slovenskej republiky - SLOV-LEX](#) >.

⁷ Spoločnosť MinnaLearn a Helsinská univerzita. *Klasifikátor najbližšieho suseda*. In *Elements of AI*. [cit. 22.2.2024] Dostupné na internete: < <https://course.elementsofai.com/sk/4/2> >.

⁸ Spoločnosť MinnaLearn a Helsinská univerzita. *Vplyvy umelej inteligencie na spoločnosť*. In *Elements of AI*. [cit. 22.2.2024] Dostupné na internete: < <https://course.elementsofai.com/sk/6/2> >.

⁹ RIVERA, J. a kolektív. *Escalation Risks from Language Models in Military and Diplomatic Decision-Making*. In *ArXiv*. [cit. 22.2.2024] Dostupné na internete: < <https://arxiv.org/abs/2401.03408> >.