

# Zsákcába kerülhet az AI a Corvinus kutatása szerint

A nagy nyelvi modellek, mint például a ChatGPT elterjedése valójában egyre csökkenti a nyilvános tudásmegosztást az online kérdezz-felelek platformokon, s ezzel megnehezítheti a jövőbeli modellek képzését – erre jutott a Budapesti Corvinus Egyetem frissen publikált tanulmánya.

A Budapesti Corvinus Egyetem nemrégiben publikált nemzetközi [tanulmánya](#) hiánypótló módon vizsgálja a ChatGPT hatását az online közösségekre, amelyek az interneten megosztott nyilvános tudáshoz járulnak hozzá, ez a hatás pedig a mesterséges intelligencia jövőjét is befolyásolja. A kutatók többek között arra mutattak rá, hogy a mára mindennapjaink szerves részévé vált nagy nyelvi modellek (LLM, large language model) potenciális adat- és tudásforrásokat helyettesítő szerepe az ember által generált tartalom csökkenéséhez vezethet. Ez jelentős gondot okozhat a jövőbeli modellek fejlesztésében, mert nem lesz elég adat a további képzésükhöz. A publikáció következtetése szerint, ha a nyelvi modellek miatt kevesebb nyílt, ember által létrehozott adat keletkezik, akkor ezek a modellek végső soron korlátozzák saját jövőbeli képzési adatforrásaikat és hatékonyságukat.

A most közzétett munka középpontjában Johannes Wachs, a Corvinus docense és nemzetközi kutatótársai a Stack Overflow nevű, kérdéseket és válaszokat tartalmazó mérvadó programozási szakmai weblap aktivitását vizsgálták a ChatGPT, az egyik legnépszerűbb LLM-modell megjelenése utáni fél évben. Az eredmények szerint a vizsgált időszakban a Stack Overflow aktivitása 25%-kal csökkent az orosz és kínai társaihoz képest, ahol a ChatGPT-hez való hozzáférés korlátozott, valamint a hasonló matematikai fórumokhoz képest, ahol a ChatGPT kevésbé hatékonyan alkalmazható. Jelentős csökkentek a felhasználók által közzétett bejegyzések is minden tapasztalati szinten, a kezdőtől a szakértőig. A ChatGPT tehát eltorlasztotta a Stack Overflow új bejegyzéseit, beleértve a kiváló minőségű tartalmakat is – vonták le a következtetést a kutatók.

## Az újrafelhasználás paradoxona

A tanulmány szerint a nagy nyelvi modellek hatására az emberek online szokásai jelentősen megváltoznak, másképp fogják keresni, létrehozni és megosztani az információkat. Ha a ChatGPT-hez hasonló LLM-ek elkezdik helyettesíteni a hagyományos keresési és lekérdezési módszereket, akkor éppen azt az emberi viselkedést szoríthatják ki, amely a modellek eredeti képzéséhez szükséges adatokat generálta. Egy ilyen, „újrafelhasználás paradoxonaként” is ismert eltolódásnak pedig jelentős társadalmi és gazdasági következményei lehetnek.

Ez a helyettesítés a nyílt web jövőjét fenyegeti, mivel a mesterségesintelligencia-modellekkel folytatott interakciók valójában nem bővítik az online tudást, a digitális közjót. Mindez gyengítheti a jövőbeli modellek képzési adatainak minőségét, mivel a gép által generált tartalom valószínűleg nem tudja teljes mértékben helyettesíteni az emberi kreativitást és éleslátást.

„Kutatásaink során arra jutottunk, hogy a ChatGPT csökkenti annak valószínűségét, hogy nyilvános webhelyeken tegyünk fel és vitassunk meg kérdéseket. Ez egyrészt hátrányos, mert mások gyakran tanulnak ezekből a beszélgetésekből, másrészt az AI nem fogja tudni kellő mennyiségű és minőségű új internetes tartalomból továbbfejleszteni magát. Ha a mesterséges intelligenciát csak mesterséges intelligencia által generált tartalommal képezzük, az olyan, mintha egy fénymásolat fénymásolatát készítenénk, és egyre kevésbé kielégítő eredményeket kapnánk” – mondta a tanulmány kapcsán Johannes Wachs, a Budapesti Corvinus Egyetem kutatója. Hozzátette: „Tudjuk, hogy azok a nyílt

internetről származó emberi visszajelzések, amelyeket a nagy nyelvi modelleknek adnak, megkönnyíthetik az LLM-ek tanulását. Azok az adatok viszont már nem lesznek közkinccsek, amelyek a magántulajdonban lévő nyelvi modellekkel való interakciókból származnak, hiszen azok az LLM-ek tulajdonosaié. Ennek érdemes a tudatában lennünk, mert jelentős következményekkel járhat mind a nyilvános internetre, mind a mesterséges intelligencia jövőjére nézve.”

## Egyre többet ér az adat tulajdonjoga

A kutatók következtetései szerint egyre fontosabb lesz gazdasági szempontból, hogy ki hozza létre az adatot és ki birtokolja. Ahogy az adatok értékesebbé válnak, egyre nagyobb lesz az érdeklődés az iránt, hogy az adatok létrehozói hogyan tudják maguknál tartani az érték egy részét. A ChatGPT-hez hasonló mesterségesintelligencia-alkalmazások tehát politikai és gazdasági nyerteseket és veszteseket fognak generálni, és hozzájárulhatnak az emberek és a cégek közötti egyenlőtlenséghez.

A [kutatás](#) - amely a PNAS Nexus szeptemberi számában jelent meg - szerzői Johannes Wachs, a Budapesti Corvinus Egyetem Adatelemzés és Informatika Intézetének egyetemi docense, R Maria del Rio-Chanona, a University College London Számítógéptudományi tanszékének adjunktusa és Nadzeya Laurentsyeva, a Münchener Lajos-Miksa Egyetem (LMU) Gazdasági Karának adjunktusa.

Sajtókapcsolat:

- [press@uni-corvinus.hu](mailto:press@uni-corvinus.hu)



© Budapesti Corvinus Egyetem  
Johannes Wachs, a Budapesti Corvinus Egyetem kutatója.

Eredeti tartalom: Budapesti Corvinus Egyetem

Továbbította: Helló Sajtó! Üzleti Sajtószolgálat

Ez a sajtóközlemény a következő linken érhető el:

<https://hellosajto.hu/15580/zsakutcaba-kerulhet-az-ai-a-corvinus-kutatasa-szerint/>