

Az értéket a szelekció és a benne lévő szaktudás adja

Hozzászólás „A szintetikus dokumentumok és a jövő archívumai” vitához

Milyen társadalmi következményekkel járhat a szintetikus dokumentumok tömegessé válása és mesterséges intelligencia általi felhasználása?

Válaszomban megfordítom a kérdések sorrendjét, az utolsót veszem előre, így logikusabb talán megágyaznom annak a gondolati halmaznak, amelyben a válaszomban mozgok. A mesterséges intelligencia társadalmi hatásait felsorolni is meghaladná a terjedelmet, ugyanez igaz a szűkebb halmazra: a generatív mesterséges intelligencia szolgáltatások hatása is nagyon sok van. Néhányat kiemelek, melyeket a jelen kontextusban lényegesnek tartok.

A szóbeli írásbeliség problémája: a digitális kultúrával együtt megjelent a szóbeli írásbeliség problémája, az írás és a beszéd összemosódik (chatkultúra, online és e-mailes kommunikáció mennyiségének növekedése), ez a generatív mesterséges intelligencia szolgáltatások szövegtermelő képességével megnövekszik, sokkal több, hosszabb, kiterjedt digitális szöveg jön létre (itt most csak azokat emelem ki, amelyeket gépek írnak emberek helyett). A generált szövegek a szóbeli írásbeliség egy új szintjét hozzák.

Az utánzás problémája: mindig is használtunk eszközöket az alkotás folyamatának segítségére, eszközhasználó faj vagyunk. Most egy olyan technológia jelent meg, ami nagy (várhatóan majd teljes) sikerrel utánozza az emberi viselkedéseket a digitális térben (később a robotokban a fizikai térben) és szövegeket. Nem egyszerűen arról van szó, hogy eszközökkel támogatjuk az alkotást, létrehozást, hanem erre a célra létrehozott tömeges szolgáltatás jelenik meg.

A kölcsönvett tudás problémája: a generatív mesterséges intelligencia saját tudás nélküli használata nemcsak forráskritikai, vagy tudatossági problémákhoz vezet el, nem is beszélve az önismeret megbicsaklásáról, hanem hosszú távon azt okozhatja, hogy a használó termékké válik. A jó stratégia a kölcsönvett tudás beépítése, az őszinte, tudatos és kíváncsi használat. A szintetikus dokumentumok megítélésénél zavaró lehet, hogy nem tudjuk, a szerző milyen tudásszinten van a témában, mennyi ebből az ember, és mennyi ebből a kölcsönvett tudás. Nem fogjuk tudni. De azt, hogy hitelesíti a kontextust, a szándékot, az időt, a lokációt, tudjuk azonosítani. Kisebb lépésekben már találkoztunk ezzel a jelenséggel, az íródeák a királyi udvarban vagy a jó tollú titkárnő egy felsővezető mellett vagy egy politikus beszédírója, hasonló helyzet volt. Itt hiúságunkat sérti, hogy esetleg ez lehet egy gép, és szakmaiságunkat zavarja, hogy léptékváltozásról van szó.

A beépülés problémája: a szintetikus dokumentumok azonosítását nyilvánvalóan megnehezíti, sőt véleményem szerint lehetetlenné teszi majd az, hogy – amellet, amikor egy chatbottal vagy MI-ügynökkel közösen hoznak létre dokumentumot

ezek beépülnek a szolgáltatásokba: automatizált helyesíráellenőrző, vagy szebb stílust adó, különböző országokba, különböző nyelveken automatikusan leveleket küldő szolgáltatások, szerződéseket előkészítő szolgáltatások stb. – millió olyan szolgáltatás van és lesz, ahol a mesterséges intelligencia (hasonlóan az internethez) elbújik. Nem mondjuk internetezésnek, amikor útvonalat tervezünk vagy levelezünk vagy chatszolgáltatáson keresztül telefonálunk. Nem is fogjuk érzékelni, amikor sok kisebb-nagyobb módosítást generatív eszközök fognak tenni.

Lehetne beszélni a dezinformációról, lehetne beszélni a tudatosságról, az ősforráshoz nyúlás igényéről, a kritikus gondolkodásról, az őszinteségről, az érzelmi torzulásokról, a digitális tér valóságképző hatásának növekedéséről, a kényelmes igazságválasztás világáról, de talán a fentiek vonatkoznak leginkább a témánkra.

Az emlékezetőrző intézmények képesek lesznek-e (képesnek kell-e lenniük) a szintetikus dokumentumok felismerésére?

Nem. A jó minőségűeket, az értő emberek által készítetteket biztosan nem, természetesen lesznek olyan minőségű vagy igényességű/igénytelenű anyagok, amelyek egyértelműek lesznek. Lehet gépi megoldásokat is használni gépi megoldások ellen, mellett – itt az látszik, hogy ezek hatékonysága a szöveg terjedelmének növekedésével növekszik, tehát csak hosszú szövegek esetében működnek jól.

Az alapprobléma véleményem szerint nem a felismerés technikai mikéntje. A valódi kérdés az, hogy eddig mit tároltunk és miért? Azért tároltunk, mert megőrizni, illetve kutathatóvá akartunk tenni valamit. Ez az alapcél most, a mesterséges intelligencia korában sem változott. De mi változott pontosan? Az, hogy a technológia miatt kezelhetetlenül több lett a szöveg, és nem tudjuk biztosan, ki a szerző, és mi járt a fejében.

Ami drasztikusan megváltozott, az az egyszer használatos szövegek mennyisége. Úgy látom, hogy a szóbeli írásbeliség világában a generált szövegek valójában csak egy újfajta szóbeliséget jelentenek. Mi az archívumban az írást, a maradandót, és nem a beszédet akarjuk rögzíteni. Csináljuk ezt tovább. A jövőbeli szelekciós döntéseim alapja tehát nem az lesz, hogy gép írta-e, hanem hogy funkcionális, eldobható szövegről van-e szó, vagy megőrzendő értékről.

Nem aggódom annyira a hallucinációk miatt, egy technológia kezdeti szakaszának a jellemzője, minél több és minél funkcionálisabb mesterséges intelligenciánk lesz, ezek lecsökkennek az emberi társadalomban megszokott szintre.

Képesnek kell lenniük az igénytelenül kivitelezett (adathiány, ellenőrzéshiány, formai inkonzekvencia stb.) szintetikus dokumentumok felismerésére, és ebben pontosan segítséget nyújthatnak majd a jól fejlesztett archiváló MI-k az emberi szak tudás mellett.

Az archívumoknak, ha felismerik, meg kell-e őrizniük a szintetikus dokumentumokat? Ha igen, hogyan?

Talán a változás az, hogy most már elsősorban nem az eredetiség/szerző, hanem a szelekció a kulcstudás. Mindent lehetetlen, és nem is kell. A kontextussal magyarázott szelekció igen izgalmas szakmai vitákat és adattárakat eredményezhet.

Ugyanúgy dokumentumokat kell őrizni az archívumoknak, kiszűrni az igényteleneket, és kontextus, felelősség, hatás, szándék, felhasználás és más változók mentén megítélve tárolni. Úgy gondolom, hogy nem annak kell változnia, hogy szintetikus-e a dokumentum, vagy sem, hanem a szelekciós szempontoknak kell összetettebbé válniuk.

De hogyan szelektáljak? A döntésnek számos tényezője van, amelyeket mérlegelni kell:

- Az időpillanat: mikor jött létre a dokumentum?
- A hitelesítő személy: ki az, aki magáénak vallja, aki a nevét adja hozzá?
- A kontextus: milyen körülmények között született?
- Egyszer használatos dokumentum?
- Van hatása? Kapcsolódik más dokumentumokhoz vagy emberekhez?

Itt merül fel a már jól ismert kérdés: a tipikust vagy a kiemelkedőt tároljuk? A múzeumi analógia szerint a hétköznapi tárgyak is válhatnak 20 év múlva fontos üzenetké a jövőnek. Ugyanez igaz a digitális térre is.

A szintetikus dokumentumok elszaporodása megváltoztatja-e majd az archívumok „hiteleshely” mivoltát a közfelfogásban?

Szerintem igen, és pozitív irányban. Ha bizalmi válság van a generált tartalmak miatt, akkor még jobban megnő az értéke annak, ha én mint archívum garantálni tudom az „ősforrást”. Úgy látom, hogy az archívum legértékesebb elemeivé az emberi beavatkozás (szelekció), a szakértői tudás és az időpecsét válnak. Nem védekezésbe kell szorulnunk, hanem felismerni, hogy a hitelességünk az új eszközökkel (például kontextuális adatok, statisztikák mentése) elképesztően pontosá válhat. Azt tudom mondani: garantálom, hogy ez a könyv vagy irat az adott időpillanatban létezett és pontosan így nézett ki. El tudom mondani a történetet, miért ezt, miért így, és miért akkor szelektáltam. Ennek ugyanolyan értéke lesz (ha nem több), mint magának a megőrzésnek.

Úgy látom, hogy átíródnak az értékrendszerek. Ha a gép segítségével bárki könnyen ír sok cikket, akkor a mennyiség értéke csökken, és más mérőszámok, például a terepmunka vagy a személyes jelenlét értéke nő meg. Hiszek abban, hogy ami technológiailag mindenki számára elérhető, annak az értéke nulla; az értéket a szelekció és a benne lévő szaktudásom adja. Nem az archívumok értéke változik meg, sőt az növekedni fog – de az üzenetük, amit maguknak magukról, illetve másoknak magukról mondanak, illetve az eddig mért ipari társadalmi értékek közül néhány nem lesz annyira fontos. Ez minden szakmában lejátszódik az MI miatt. Jönnek új, fontos értékek ezek mellé.

Rab Árpád információs társadalom kutató, kutatásai során a digitális technológiák és a társadalom egymásra hatását vizsgálja. Elméleti kutatásai és oktatói munkássága mellett konkrét fejlesztéseket vezet (jelenleg mesterséges intelligencia és humanoid robotok területén), üzleti tanácsadást végez, illetve a tudománykommunikáció elkötelezett híve, számos konferencián és ismeretterjesztő műsorban szólal meg.

Így hivatkozzon erre a cikkre:

<https://dx.doi.org/10.22503/inftars.XXVI.2026.1.8>