



MI Újság

A Nemzeti Közsolgálati Egyetem Információs Társadalom Kutatóintézete havi hírlevele a mesterséges intelligencia alkalmazásáról, társadalmi hatásairól és kérdéseiről

2024 május

Az NKE ITKI honlapja: itki.uni-nke.hu

A hírlevél tartalma a Creative Commons Nevezd meg! – Ne add el! – Így add tovább! 4.0 Nemzetközi Licenc feltételeinek megfelelően használható.



**NEMZETI
KÖZSZOLGÁLATI
EGYETEM**
LUDOVIKA

TARTALOMJEGYZÉK

Etika és jog

- A jogi modellek szinte minden esetben hallucinálnak
- Az MI-főhadiszállások aranykora Indiában

Trendek

- Nemzetközi tudományos jelentés a fejlett mesterséges intelligencia rendszerek biztonságáról
- Ne higgyünk az MI-rendszereket övező túlzó véleményeknek!
- A kis nyelvi modellek felfutása, avagy senki sem szereti a mindenes szerszámokat

Működésben

- A digitális túlvilág felfedezése
- Svájc is beszállna az MI-versenyfutásba
- A mesterséges intelligencia hipnotikus irodalmi álmai: így születhet egy új műfaj





Etika és jog

A jogi modellek szinte minden esetben hallucinálnak

Az általános célú nagy nyelvi modellek kibontakozása a túlzott várakozások és a sűrűn előforduló hibás működések közötti erőterben történik. Az ismert, hogy ezek a rendszerek strukturális okokból hajlamosak válaszaikban megbízhatatlan, kitalált információkat, úgynevezett hallucinációkat adni. Ennek a problémának a megoldására alkalmazzák a kereséssel támogatott szöveggenerálást (retrieval-augmented generation, RAG). Ebben az esetben egy adott szakterület specializált kérdésmegválaszoló rendszerét (chatbot) úgy alakítják ki, hogy az adott szakterület megbízható adatbázisából származó információkkal egészítik ki a nagy nyelvi modell működését. Különösen nagy jelentősége van ennek a jog területén, ahol ezektől a rendszerektől azt várják, hogy pontosabb és megbízhatóbb jogi információkat nyújtsanak egy nagy nyelvi modell és egy jogi dokumentumokat tartalmazó adatbázis integrálásával. A nagy szolgáltatók azt ígérik, hogy szolgáltatásaik mentesek a hallucinációktól. Bizonyítékot azonban nem szolgáltatnak állításaik alátámasztására és még a „hallucináció” fogalmát sem határozták meg pontosan, így nehéz megítélni a szolgáltatásaik valós megbízhatóságát. A Stanford Egyetem kutatói a Stanford Institute for Human-Centered Artificial Intelligence vezetésével két vezető szolgáltató (LexisNexis, Thomson Reuters/Westlaw) állításait tették egzakt vizsgálat tárgyává. A kutatás szerint a két cég eszközei valóban csökkentik a hibákat az olyan általános célú MI-modellekhez képest, mint a GPT-4. Ez jelentős javulást jelent, de még ezek a teszteszabott jogi MI-eszközök is riasztóan sokszor hallucinálnak: az esetek több mint 17%-ában – minden hatodik lekérdezésből egy esetben – hibás információt produkáltak.

[AI on Trial: Legal Models Hallucinate in 1 out of 6 Queries](#)

Az MI-főhadiszállások aranykora Indiában

Idén rekord számú országban kerül sor parlamenti vagy egyéb választásokra – olyan országtömörülések állampolgárai is a szavazó urnákhoz járultak, mint a sok szempontból meghatározó szereplőnek tartott Európai Unió. Az egyik megkülönböztetett figyelemmel kísért szavazás Indiáé, amely nem csupán a világ egyik (becslések szerint első, vagy második) legnépesebb országa, hanem a technológiai fejlesztéseknek, az informatikának, kommunikációs technológiáknak és a mesterséges intelligencia technológiáknak is az egyik erőközpontja, Ázsiában biztosan. Az idei évben rendezett és valóban sokak szerint ott is sordöntőnek prognosztizált választási küzdelemben is megtalálta helyét az új MI-technológia: iparági becslések szerint a helyi pártok az idei kampányban több mint 50 millió dollárt költöttek MI-vel generált kampányanyagokra. A sokat emlegetett deepfake technológia különösen kedvelt eljárás – természetesen etikus alkalmazásban, azaz nem az ellenfelek lejáratására, hanem az adott jelölt megjelenítésének javítására. Gombamód szaporodnak az országban az

ún. AI War Room-ok, azaz MI-főhadiszállások, ahol egy-egy párt (de inkább egy-egy konkrét jelölt) MI-megjelenéseit tervezik, generálják, szolgáltatják. A soknyelvű szövetségi államban egyébként az MI-technológiával készülő választási tartalmaknak a pusztán elérésszám növelésében is hatalmas szerep juthat. Diggaj Mogra, a Jarvis Consulting, egy mumbai székhelyű politikai tanácsadó cég igazgatója szerint azonban a mesterséges intelligencia még gyerekcipőben jár. Nincs akkora hatása, mint a helyszíni kampányolásnak, mert jelenleg csak „támogató szerepet játszik”. A következő, 2029-es választásokon az MI már fontos szerepet fog betölteni, de jelenleg még csak kísérleti fázisban van.

[The booming business of AI war rooms during India's elections](#)





Trendek

Nemzetközi tudományos jelentés a fejlett mesterséges intelligencia rendszerek biztonságáról

Az Egyesült Királyság tudományos kutatásért és iparpolitikáért felelős minisztériuma az újonnan létesített MI-biztonsági intézettel és egy nemzetközi szerzői gárdával együttműködésben megjelentette a fejlett mesterséges intelligencia biztonságáról szóló jelentését. Az időközi beszámoló célja, hogy az Egyesült Királyság és Dél-Korea által 2024. május 21-22-én közösen megrendezett AI Seul Summit számára vezérfonalul szolgáljon. Ez az első naprakész, nemzetközi tudományos jelentés a fejlett mesterséges intelligencia rendszerek biztonsági problémáiról. Bemutatja a jelenlegi és a várható MI-rendszerek képességeit, a várható kockázatokat, valamint a közpolitikai lehetőségek jobb megismerése érdekében a kockázatok mérséklésére és értékelésére szolgáló megközelítéseket. A jelentés az általános célú MI-re összpontosít, és egy „mesterséges intelligenciamodellt akkor tekint általános célúnak, ha képes a feladatok széles körének elvégzésére, vagy adaptálható ezekhez”. Bár a 30 szakértő között bizonyos fokú konszenzus alakult ki a fejlett MI-rendszerek képességeivel és kockázataival kapcsolatban - különösen a jövőben várhatóan kifejlesztendő rendszerek esetében - jelentős nézeteltérések mutatkoztak. Ennek az az oka, hogy egyes szakértők szerint a fejlődés lassulása a legvalószínűbb, míg más szakértők szerint rendkívül gyors fejlődés lehetséges. A tudományos konszenzus hiánya miatt a jelentés inkább csak a tudományos ismeretek jelenlegi állapotának áttekintésére szolgál. A jelentés nem tesz politikai vagy szabályozási ajánlásokat, a fejlett MI-rendszerek biztonsága körüli „kód tisztázásával” arra törekszik, hogy megkönnyítse a döntésekről szóló konstruktív vitákat. A jelentés következtetése az, hogy az általános célú MI jövője rendkívül bizonytalan, és a társadalmi és kormányzati döntésektől függ. A készítőik folyamatos nemzetközi együttműködésre szólítanak fel az MI-technológiák biztonságos és előnyös fejlesztésének biztosítása érdekében.

[International Scientific Report on the Safety of Advanced AI](#)

Ne higgyünk az MI-rendszereket övező túlzó véleményeknek!

Az MI-rendszerek fejlődését kísérő felfokozott várakozásokat számos túlzó előrejelzés kíséri. Ezek a termelékenység, a nemzeti jövedelem nagyon jelentős növekedését jelzik már rövid távon is. Daron Acemoglu, a MIT amerikai egyetem vezető kutatója és a munka jövőjét firtató kutatási projekt vezetője cikkében a mesterséges intelligencia gazdasági-társadalmi hatásainak realistább megítélését szorgalmazza. Számításai szerint az MI-rendszerek hatására a teljes tényezőtermelékenység (total factor production, TFP) tíz év alatt csak 0,66% százalékkal fog nőni, ami 0,06 százalékos éves növekedést jelent. Mivel az MI-rendszerek alkalmazását befektetési fellendülés is kíséri, a GDP növekedése ennél nagyobb, talán 1-1,5%-os lehet. Ezek a számok sokkal kisebbek, mint az ismert üzleti tanácsadó cégek nagy nyilvánosságot kapott adatai. Ha nagyobb számokat akarunk elérni, akkor vagy mikroszinten kell fokozni a termelékenység növekedését, vagy fel kell tételeznünk, hogy a gazdaságban sokkal több feladatot érinthet az MI-rendszerek alkalmazása. Hasonlóképpen, nem valószínű, hogy a feladatok 4,6%-ánál sokkal több feladatot vesznek át majd az MI-rendszerek, mivel az MI közel sem képes a legtöbb manuális vagy szociális feladat elvégzésére (beleértve az olyan látszólag egyszerűnek tűnő, bizonyos szociális aspektusokkal rendelkező funkciókat, mint például a könyvelés). Ha az elmúlt másfél évben fel is gyorsultak az MI beruházások, még hosszú út áll előttünk, mire az MI széles körben elterjed. A közgazdasági elmélet és a rendelkezésre álló adatok a mesterséges intelligenciával kapcsolatban szerényebb, reálisabb kilátásokat indokolnak. Ha kritikátlanul átvesszük a techno-optimizmust vagy hagyjuk, hogy a technológiai ipar határozza meg a teendőket, akkor az MI nyújtotta lehetőségek nagy részét elszalaszthatjuk.

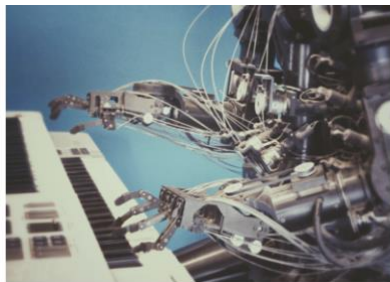
[Don't Believe the AI Hype](#)

A kis nyelvi modellek felfutása, avagy senki sem szereti a mindenes szerszámokat

Még tavaly – a nagy nyelvi modellek mámorában – alkotta meg a Microsoft cég a „kis nyelvi modell” kifejezést. Lényegében az első Phi modelljüket bemutató „Textbooks Are All You Need” című tanulmányukban használták először ezt a formulát, és azóta az SLM-iparág valóságos szárnyalásnak indult. Sokan arról beszélnek, hogy könnyen lehetséges, hogy ezek a Small Language Model-ek lesznek a generatív MI-technológiák következő nagy sztárjai. Az SLM-ek mellett szóló érvek elég egyértelműek. Az óriási méretű alapmodellek generalista alkalmazásokra valók, és miközben továbbra is rendkívül drága az üzemeltetésük, rengeteget hallucinálnak, valamint biztonsági sebezhetőségekkel és megbízhatósági problémákkal küzdenek, amikor domain-specifikus környezetben alkalmazzák őket. A specializált alkalmazásra más eszközök valók, amelyek biztonságosabbak és megfizethetőbbek. A közelmúltban két új szereplővel is bővült ez az ígéretes SLM-piac: a Microsoft előrukkolt az új Phi-3 modell családjával, amely rendszeresen felülmúlja sokkal nagyobb vetélytársainak képességeit. Másrészt pedig az Apple hozzáférhetővé tette a forráskódját az OpenELM névre keresztelt, mobil környezetre optimalizált kis nyelvi modell családjának. A nagy alapmodellek tehát továbbra is uralják a generatív mesterséges intelligencia területét, és ez így is marad, amíg a skálázási törvények érvényesülnek. Az kis nyelvi modellek azonban minden bizonnyal a piac egy fontos szegmensét fogják elfoglalni a jövőben.

[Nobody Likes a Know-It-All: Smaller LLMs are Gaining Momentum](#)





Működésben

A digitális túlvilág felfedezése

A Synthetic Minds podcast legújabb epizódjának vendége Richard K. Morgan, az „Altered Carbon” című könyv vizionárius szerzője volt. Az adás házigazdája Dr. Mark van Rijmenam volt, aki stratégiai jövőkutatóként és innovációs szakértőként határozza meg önmagát. A Synthetic Minds podcastban a tudományos-fantasztikum, illetve a valós technológia és a társadalmi változások találkozási pontjairól szoktak beszélgetni a résztvevők. A beszélgetés apropóját Richard K. Morgan regénye adta, amelyből népszerű Netflix-sorozat is készült. A könyv egy olyan jövőt mutat be, ahol az emberi tudat digitalizálható és testek között átvihető. Ez az elképzelés nemcsak a személyes identitást és a halandóságot definiálja újra, hanem mély etikai kérdéseket is felvet az ilyen technológia társadalmi és egyéni következményeivel kapcsolatban. A történet középpontjában a technológiai és az emberi természet sötétebb aspektusai állnak, azaz a digitalizált, testtől független tudat által meghatározott világ társadalmi és egyéni problémái. A beszélgetésnek azonban vannak gyakorlati irányultságú részei is. A beszélgetők megvitatták, hogy a spekulatív gondolkodás és történet hogyan segíthet a vezetőknek a stratégiai előrelátás fejlesztésében. A jövőbeli forgatókönyvek elképzelésével a vállalkozások jobban fel tudnak készülni a technológiai fejlődésre és hatékonyabban tudnak alkalmazkodni ahhoz. Ennek a munkának részei a trendelemzés, a forgatókönyv-tervezés és az innovációs műhelymunkák, melyek a proaktív megközelítést támogatják, a reaktív helyett. A regény végső soron felhívásként szolgál, hogy mérlegeljük az innovációk szélesebb körű következményeit, valamint hogy a technológiai fejlődés hogyan alakíthatja át a társadalmi normákat és az egyéni jogokat.

[The Human Element: Exploring the Digital Afterlife with Richerd K. Morgan](#)

Svájc is beszállna az MI-versenyfutásba

Svájcról sok minden eszébe jut az embernek, a finom tejcsokoládétól a legendásan pontos villamosított vasúthálózatokig, valamint a pontosság szinonimájaként emlegetett svájci óraszerkezetekig, de a csúcstechnológiát ritkán kapcsoljuk ehhez a festői alpesi országhoz. És ezt bizony rosszul tesszük. Bármilyen meglepő is, Svájc a 21. századi mesterséges intelligencia-korszak egyik nagyhatalmává nőheti ki magát. Figyelemre méltó, hogy a Tortoise Global AI Index legutóbbi, 2023-as összeállítása összességében a kilencedik (!) helyre teszi a svájci MI-ökoszisztéma teljesítményét. Az adatsorokat megnézve rögtön kiténik az is, hogy Svájc alapvetően három területen teljesít kiemelkedően jól. Nagyon erős az MI-hez kapcsolódó kutatás-fejlesztésben, ahol globálisan a negyedik helyen áll; a technológiák ipari bevezetésében, és a szakemberképzésben, ahol a kilencedik helyen áll. A svájci generatív MI-piac nagyságát 650 millió dollárra becsülik az idei év végéig, a tágabb értelemben vett mesterséges intelligencia technológiák piaca pedig az előrejelzések szerint elérheti az 1,74 milliárd dollárt is az év végére. Bár Svájc kiemelkedő fontosságú szereplő a mesterséges intelligencia fejlesztések területén, a technológia megvalósításában még nem elég kiegyensúlyozott. Számos magas színvonalú svájci MI-vállalat létezik, de a legtöbbjük inkább néhány kiválasztott részterületre (elsősorban a robotikára) összpontosít, minthogy egyenletesen oszlanának el a különböző ágazatokban. A közsféra még kevésbé változatos, a legtöbb kormányzati kezdeményezés még mindig kezdeti stádiumban van. Azzal, hogy Svájc innovációbarát környezetébe egyre több variabilitás kerül, amelyet új szereplők és különböző méretű vállalkozások is erősítenek, a nemzet abszolút MI-kapacitása is várhatóan növekedni fog.

[Swiss Made: The Alpine nation's affair with AI](#)

A mesterséges intelligencia hipnotikus irodalmi álmai: így születhet egy új műfaj

A teljes egészében a szerző által elkészített, kiadott és forgalmazott szépirodalom már jól bejártott, közismert iparág. Ugyanakkor bárki, aki csak egyszer belevágott ebbe a műfajba, az tudja: ez az egyik legnehezebb vállalkozás. A könyvpiacokat ugyanis azok a roppant marketing költségvetéssel megtámogatott könyvek uralják, amelyeket jelentős cégek forgalmaznak. A magánkiadásban megjelenő munkák esetében a leggyakoribb terjesztési marketing forma a közösségi média felületein folyó és minden szempontból „civil”, azaz nem túl professzionális reklámkampány, aminek eredményessége általában rendkívül csekély. Ezt a helyzetet kívánja orvosolni egy új generatív MI-eszköz, a Hypnovels, amely egy rövid, de nagyon látványos, hang és mozgókép effektusokból felépülő bemutató részletet készít a forgalmazandó műből. A szerző feltölt egy kiválasztott részletet, esetleg néhány háttér információt, és a mesterséges intelligencia elkészíti belőlük a bemutató anyagot. Ez kerülhet aztán a közösségi platformokra, a szokványos „Pillants bele az új könyvembe!” típusú reklámmondatokhoz képest sokkal hatékonyabb marketing eszközként. A Hypnovels a jelenlegi formájában még fejlesztés alatt áll, és azok a szerzők, akik már most is használják, segítenek majd az eszköz finomhangolásában. Ez az oka annak is, hogy a Hypnovels egyelőre még ingyenesen használható, bár a működési költségek miatt valószínűleg előbb-utóbb olyan szolgáltatássá kell majd válnia, amelyért a szerzőknek fizetniük kell.

[Hypnovel – This Ai turns novels into hypnotic dreams](#)

