

MI Újság

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

2024 február

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

- Védőkoriátok az emberi döntések irányításához az MI korában
- Az MI-rendszerek etikus katonai alkalmazása az ASIMOV program segítségével
- Széleskörű biológiai támadások elősegítésére használható MI-rendszerek kockázatai

Trendek

- A nyílt alapmodellek társadalmi hatásai
- Amerikai iskolák MI-eszközökkel „kémkednek” az öngyilkosság gondolatával játszó diákok után
- A svéd e-kereskedő óriás, a Klarna már döntött: a chatbotok mellett nincs szüksége több munkaerőre

Működésben

- Az MI és a zöld szemlélet hiánya: az új technológia igazi árnyoldala
- A mesterséges intelligencia alkalmazása az oktatásban és a tanulói teljesítmények értékelésében
- MI a rendőrségi munkában: újabb etikai kérdések, ezúttal a papírmunka támogatásában
- A (mesterséges) intelligencia paradoxona: az MI irracionálisabbá teheti a piacokat





Etika és jog

Védőkoriátok az emberi döntések irányításához az MI korában

Urs Gasser és Viktor Mayer-Schönberger frissen megjelent könyvükben azt mutatják be, hogy a technológiai megoldások gyors elfogadása hogyan vezethet olyan eredményekhez, amelyeket nem mindig akarunk. A szerzők elmagyarázzák, hogy maga a társadalom hogyan biztosíthat a digitális korszakhoz jobban illeszkedő „védőkoriátokat”, amelyek lehetővé teszik az egyéni választást, miközben figyelembe veszik a társadalmi jót, ösztönzik a rugalmasságot a változó körülmények közepette, és végül segítenek abban, hogy jobb döntéseket hozzunk korunk legijesztőbb problémáinak - például a globális igazságtalanságnak és az éghajlatváltozásnak - kezelésében. Ahogy haladunk előre a digitális korban, úgy változik a tényleges hatalom helye és így a szabályozás fókuszja is. Korábban a hálózat volt a meghatározó terület: Ki és milyen összeköttetéssel rendelkezik? Ki ellenőrzi a hálózat szűk keresztmetszeteit? Milyen interakciókat enged meg a hálózat, és melyeket korlátoz? Milyen virtuális terek és platformok alakulnak ki? Újabban az adatokra összpontosítottunk: Ki gyűjthet és használhat fel adatokat? Milyen mintákat tárnak fel a hatalmas adatmennyiségek, amelyeket a kisebb adathalmazok elrejtenek, és ki láthatja ezeket a felismeréseket? Mit tanulnak majd a gépek az adatokból? Ennek megfelelően pedig az adatkormányzás és az információ geopolitikája vált napjaink fontos témájává. A könyv szerzői szerint azonban a mesterséges intelligencia korában az adatokra való összpontosítás téves, vagy legalábbis veszélyesen hiányos. Az MI valódi ígérete ugyanis nem az ismeretekhez való jobb hozzáférés, hanem a jobb döntéshozatal. Minél inkább felismerjük az emberek kognitív hibáit, és minél inkább megtapasztaljuk az adatvezérelt gépi tanulás erejét, annál nagyobb a kísértés, hogy a döntési jogköröket a gépre bizzuk. A hálózat- vagy az adatkormányzás helyett a döntés-irányításra (döntés-kormányzásra) kell összpontosítanunk, erre a fókuszváltásra azonban nem vagyunk felkészülve. A könyv szerint a valódi innováció, amelyre szükségünk van ahhoz, hogy ez megvalósuljon, kevésbé technikai, mint inkább társadalmi jellegű.

[Guardrails: Guiding Human Decisions in the age of AI](#)

Az MI-rendszerek etikus katonai alkalmazása az ASIMOV program segítségével

Az Egyesült Államok Védelmi Minisztériumának kutatásokért felelős részlege, a Defense Advanced Research Projects Agency (DARPA) kutatási felhívást hirdetett,

melyben innovatív javaslatokat vár az autonóm rendszerekre vonatkozó referenciaértékek kutatására és fejlesztésére. A kutatásoknak olyan innovatív megközelítéseket kell vizsgálniuk, amelyek forradalmi előrelépést jelenthetnek a tudomány, az alkalmazott eszközök vagy a rendszerek terén. Az ASIMOV (Autonomy Standards and Ideals with Military Operational Values) program az amerikai hadsereg kísérleti programja, amelynek célja az autonóm rendszerek használatával kapcsolatos etikai problémák kezelése, valamint egy olyan objektív és kvantitatív keret létrehozása, amely alkalmas a jövőbeli autonóm felhasználási esetek etikai nehézségeinek és az autonóm rendszerek felkészültségének mérésére és értékelésére. Az ASIMOV program célja továbbá, hogy megteremtse az etikus autonómia nyelvezetét, amely lehetővé teszi a fejlesztési tesztelés/üzemeltetési tesztelés (DT/OT) szakmai közösségének számára, hogy értelmesen értékelje a konkrét katonai alkalmazási forgatókönyvek etikai nehézségeit és az autonóm rendszerek azon képességét, hogy etikusan teljesítsenek e forgatókönyveken belül. Az autonóm és MI-technológiák gyors fejlődése és közelgő elterjedése mind a polgári, mind a katonai alkalmazásokban megbízható és kvantitatív keretet igényel, nemcsak az autonóm rendszerek műszaki, hanem – ami talán még fontosabb – az etikai képességének mérésére és értékelésére, amint azok a kutatás-fejlesztésen túl is megjelennek.

[Autonomy Standards and Ideals with Military Operational Values \(ASIMOV\)](#)

Széleskörű biológiai támadások elősegítésére használható MI-rendszerek kockázatai

Az mesterséges intelligencia rendszerek gyors fejlődése nyomán felvetődött az az aggasztó gondolat, hogy elterjedésük a biológiai fegyverek kifejlesztését is jelentősen megkönnyítheti. Az MI-rendszereknek ez a potenciális alkalmazása azért különösen aggályos, mivel azok nem állami szereplők és magánszemélyek számára is hozzáférhetőek, emellett az MI technológiák ilyen rohamos fejlődésével nem tud lépést tartani a kormányzatok szabályozási tevékenysége sem. Az amerikai RAND Corporation szakértői a lehetséges kockázatokat egy olyan szimuláció keretében vizsgálták, ahol szakértői csoportok azzal voltak megbízva, hogy meghatározott forgatókönyv alapján a rosszindulatú csoportok szerepét játszva (red teaming) megtervezzenek egy biológiai támadást. Egyes csoportoknak csak az internethez volt hozzáférésük, másoknak az internethez és a nagy nyelvi modellekhez (LLM) is. A több nagy nyelvi modellt érintő kutatás végső megállapítása az, hogy a biológiai fegyverekkel végrehajtott támadás tervezése jelenleg meghaladja a nagy nyelvi modellek képességeinek határát. A szerzők nem találtak statisztikailag szignifikáns különbséget a nagy nyelvi modellek segítségével vagy anélkül készített tervek életképességében. Bár azonosíthatók voltak a nagy nyelvi modellek aggasztó válaszai is, ezek általában az interneten könnyen elérhető információkat tükrözték, ami arra utal, hogy az LLM-ek nem növelik jelentősen a biológiai fegyverek támadásának tervezésével kapcsolatos kockázatokat. A kutatók a későbbi kutatások során növelnék

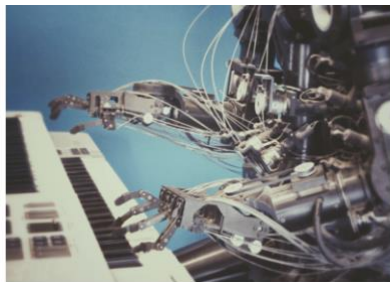
a vizsgált nagy nyelvi modellek számát, több kutatót vonnának be, és javítanák a tesztelési folyamat minőségét.

[The Operational Risks of AI in Large-Scale Biological Attacks](#)



Trendek





Működésben

Az MI és a zöld szemlélet hiánya: az új technológia igazi árnyoldala

A mesterséges intelligencia rendszerek társadalmi ismertségének berobbanásával párhuzamosan erősödtek fel a szakmai, sőt a politikai és laikus közbeszédben is a szinte minden területet, társadalmi működésünket átszövő új technológiák lehetséges veszélyei, fenyegetései. A legnagyobb figyelem talán az MI-hez kapcsolódó etikai problémákat kíséri. A közvetlen kérdéstől, hogy elveszi-e az MI a valódi emberek munkáját, a gyilkos robotok témáján át egészen az emberi faj másodrendű pozícióba kényszerülésének rémképéig nagyon súlyos dilemmákat fogalmaznak meg az MI-kritikusok. Van azonban egy téma, amely gyakorlatilag teljesen kiszorul a társadalmi diskurzusnak ebből a frontvonalából: az MI-technológiák és a környezetvédelem, a fenntartható gazdaság, a „zöld társadalmak” kapcsolódása. A The New Republic szerzője nemrég megjelent cikkében annak a véleményének ad hangot, miszerint a mesterséges intelligencia a víz- és energiaigényével, a működése során keletkező e-hulladékkal és a kritikus nyersanyag készleteket felélő természetével veszélybe sodorja az emberiség fenntartható jövőjét. A szerző álláspontja szerint a mesterséges intelligencia teljesen feleslegesen veszélyezteti a Földön való kényelmes élethez szükséges fizikai erőforrásokat, és az energetikai átmenetet. Sok más környezetkárosító iparágat érthető módon nehéz megváltoztatni vagy helyettesíteni, mert az emberi civilizáció számára értéket képviselnek - ilyen például a mezőgazdaság, amely nélkülözhetetlen a mindennapi élethez. Az MI azonban nem éri meg a jelentős költségeket és áldozatokat, mivel – a szerző nézete szerint – semmi olyat nem tud nyújtani, amire valóban szükségünk van.

[The Scariest Part about Artificial Intelligence](#)

A mesterséges intelligencia alkalmazása az oktatásban és a tanulói teljesítmények értékelésében

Az angol Parlament tudományos tanácsadó intézménye, a Parliamentary Office of Science and Technology hasznos összefoglalót készített a parlamenti képviselők számára az MI-rendszerek oktatási alkalmazásáról – felsorolva a mellette és az ellene felhozott érveket, valamint ismertetve néhány fontosabb kormányzati döntést is. A generatív MI-rendszerek gyors fejlődése válaszút elé állítja az oktatást, ugyanakkor megteremti a lehetőségét annak, hogy a tanulás a megszokottól eltérő, új módokon történjen és segítséget nyújtson a pedagógusoknak a tanórák megtervezésében, a hallgatók teljesítményének értékelésében és más feladatokban is. Azonban jelenleg az MI-rendszerek oktatásban történő alkalmazásának még korai és kísérleti fázisában vagyunk, és ezt a folyamatot, a lehetséges előnyöket és a hátrányokat rengeteg bizonytalanság övezi. Többen aggodalmuknak adtak hangot azzal kapcsolatban, hogy az MI-rendszerekre való túlzott támaszkodás csökkentheti a pedagógus-tanuló kapcsolat minőségét. Aggályok merültek fel továbbá az MI-nek a tanulók íráskészségére és kritikai gondolkodására gyakorolt esetleges negatív hatásaival kapcsolatban is. Az angol oktatási minisztérium ezzel együtt 2023 novemberében jelentést tett közzé a generatív mesterséges intelligencia oktatásban való alkalmazásáról, a kormány pedig bejelentette, hogy akár 2 millió fontot is fog fordítani arra, hogy a tanárok szabadon használhassanak egy MI-alapú tananyagtervező platformot. Az MI sikeres alkalmazásához az oktatásban a következőkre lesz szükség: bizonyíték arra, hogy a mesterséges intelligencia hol lehet hatékony az oktatási eredmények elérésében; képzés és iránymutatás a pedagógusok számára; egyértelmű jogi keretek, amelyek biztosítják az MI-rendszerek oktatási alkalmazásának megfelelését az adatvédelmi követelményeknek; gondoskodni kell arról, hogy az MI-eszközök használata nem súlyosbíthatja a társadalmi egyenlőtlenségeket.

[Use of artificial intelligence in education delivery and assessment](#)

MI a rendőrségi munkában: újabb etikai kérdések, ezúttal a papírmunka támogatásában

Több vállalat és egykori rendőrök is azt vizsgálják, hogyan lehetne a generatív mesterséges intelligencia segítségével csökkenteni a rendőrségi jelentések írásával töltött időt, amelyet aztán a bűncselekmények jobb felderítésére lehetne fordítani. A louisianai székhelyű 365 Labs elnevezésű, bűnüldöző szervek számára szoftvereket fejlesztő cég és a PoliceReports.ai, amelyet egy volt floridai rendőr alapított, most olyan technológiával rukkoltak elő, amely másodpercek alatt képes rendőrségi jelentéseket írni. Ezek a fejlesztések különböző módon, de ugyanazon az elven működnek: a rendőrök megoszthatják egy incidens részleteit egy chatbottal, és egy jogilag elfogadható, MI által írt dokumentumot kapnak vissza. A kormányzati szervek és a tudósok azonban máris riadót fújnak - rámutatva a technológia két gyenge pontjára: a hibázásra, sőt a tények meghamisítására való hajlamra; és az érzékeny információk kiszivárgásának kockázatára. Az olyan nagy nyelvi modellek, mint a ChatGPT, közismertek a pontatlanságukról, és ezek a hibák - bár számos kontextusban ártalmatlanok – a bűnüldözésben való alkalmazásuk esetén életbe vágó

következményekkel járhatnak. Bár nincsenek szövetségi előírások arra vonatkozóan, hogy a kormányzati szervek hogyan használják a mesterséges intelligenciát, számos javasolt irányelv hangsúlyozza az emberi felülvizsgálat szükségességét az MI közszférában való használatára vonatkozóan. Jonathan Parham, a New Jersey állambeli Rahway egykori rendőrségi igazgatója szerint az MI által írt jelentések nem tartalmazzák az olyan árnyalt részleteket, amelyeket csak egy ember tud leírni, és hogy akár egy apró hiba vagy kihagyás is érvényteleníthetné az egész jelentést. Ezért azt javasolja, hogy a generatív MI-t inkább segédeszközként alkalmazzák, ennek megfelelően létre is hozott egy ChatGPT botot, amelyet "rendőrségi jelentés-író asszisztensnek" nevezett el.

[Experts cold on police AI](#)

A (mesterséges) intelligencia paradoxona: az MI irracionálisabbá teheti a piacokat

A világ talán legjelentősebb pénzügyi lapja, a Wall Street Journal egy rendkívül izgalmas kérdést, valamint egy nagyon érdekes és ígéretes MI-alkalmazási területet vizsgált meg egy kicsit közelebbről. A mesterséges intelligencia elemző ereje, vagyis az a képessége, hogy múltbeli adattömegekkel dolgozva az emberi szem számára nehezen észlelhető mintázatokat, rendszerességeket, netán törvényszerű ismétlődéseket fedezzen fel, különösen ígéretes eszközzé teszi ezt a technológiát a pénzügyi világ – sokszor éppen jövőbeli eshetőségeire tippelő – szakemberei számára. A technológia felgyorsuló fejlődésével szakértői körökben immár felmerült a kérdés: lehet, hogy hamarosan beköszönt az az időszak, amikor a hús-vér befektetési portfólió kezelők menthetetlenül „elavulttá” válnak egyszerűen azért, mert az algoritmusok következetesen jobb előrejelzéseket adnak majd náluk? Vajon elképzelhető-e, hogy a piacok, a tőkés gazdaságok legfontosabb gazdasági szinterei tökéletessé válnak? Olyan végső egyensúlyi mechanizmussá, ahol az árak az emberi beavatkozás különféle „fénytöréseitől” mentesen, a gazdasági realitás hű visszatükrözőiként funkcionálhatnak és ahol többé nem torzítja a valóságot az emberi beavatkozás. A cikk közgazdász szerzője rövid, de a laikusok számára is követhető módon, gazdaságtörténeti hivatkozásokkal is megtámogatva viszi végig következtetését: miközben az MI ténylegesen csökkentheti vagy akár teljesen el is tüntetheti az emberi irracionalitásból fakadó piactorzító hatásokat, a piac strukturális irracionalitásait – amelyek a piaci szereplőkre nehezedő kényszerekből erednek – nem tünteti el, sőt, akár még fel is erősítheti azokat.

[The Intelligence Paradox: AI May Make Markets Less Rational](#)