

# MI Újság

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

2024 szeptember

Az NKE ITKI honlapja: [itki.uni-nke.hu](https://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

- Elkészült az ENSZ Felső Szintű Tanácsadói Testületének jelentése az MI szabályozásáról
- Első ízben engedélyezték, hogy egy MI-chatbot befektetési tanácsokat adjon
- Hogyan győzzünk a generatív MI chatbotokkal szemben?
- Hagyjuk csak, hogy a sztochasztikus papagájok kedvükre rikácsoljanak a szaksajtóban

## Trendek

- Személyiség-bizonylatok ember mivoltunk online igazolására
- A mesterséges intelligenciával segített plagizálás problematikája: trendek és megoldások
- A generatív MI-rendszerek elszennyezik az emberi tudást megtestesítő adatokat
- Az MI-rendszerekben testet öltő innováció lenyűgöző, de még nem emelte a termelékenységet

## Működésben

- „A ChatGPT miatt hagytam ott a tanítást”
- A nyelvi modellek saját valóságérzékelést fejlesztenek ki
- Forradalmi áttörést hozhat az infarktus korai előrejelzésében egy MI
- A Google MI-szoftvere fogja segíteni a munkanélküli juttatások elbírálását Arizonában





## Etika és jog

### Elkészült az ENSZ Felső Szintű Tanácsadói Testületének jelentése az MI szabályozásáról

Elkészült és a nyilvánosság elé került az Egyesült Nemzetek Szervezete szakosított tanácsadói testülete által kidolgozott, és a sokak által az MI-kérdéskör globális szintű, multilaterális szabályozásának egyik legfontosabb lépcsőfokának tekintett jelentés. "A mesterséges intelligencia irányítása az emberiségért" (Governing AI for Humanity) címet viselő dokumentum egy „globálisan befogadó” MI-igazgatási rendszer kialakítására tesz javaslatot, amely garanciát jelenthet a jelenleg a technológia mozgatórugójaként szereplő piaci mechanizmusok korlátok közé szorítására, valamint az üzleti indíttatások nyomán nemzeti szinteken szétaprózott cselekvések nemzetközi összehangolására. A geostratégiai világ több jelentős személyiségét is magába foglaló UN High-Level Advisory Body on Artificial Intelligence most publikált szakmai anyagában hét jelentősebb ajánlást fogalmaz meg a Világszervezet, illetve az abban résztvevő nemzetállamok képviselői számára.

1. Egy olyan nemzetközi tudományos testület felállítása, amely éves jelentésekben foglalná össze az MI jelentette kockázatokat és lehetőségeket, valamint kutatásokat végezne arról, hogy az MI hogyan segítheti az ENSZ Fenntartható Fejlődési Céljainak elérését;
2. Egy olyan, évente kétszer ülésező szakpolitikai fórum felállítása, amely az MI-kormányzásra vonatkozó nemzeti törekvések „interoperabilitását” előmozdító szerveződésként segítene a jó gyakorlatok kormányok közötti terjesztésében;
3. Egy nemzetközi „MI Szabvány Tőzsde” (AI Standards Exchange) felállítása, amely előmozdítaná az MI-témakörhöz tartozó közös fogalmi keretek, szabványok, mérő- és értékelő rendszerek kialakítását;
4. Kapacitásfejlesztő hálózat kialakítása a kutatói, vállalkozói közegek szakmai felkészültségnek (tudásának, adatokkal és számítási kapacitásokkal való ellátottságának) előmozdítására;
5. Közös, globális „MI-fejlesztési Alap” felállítása: az MI-fejlesztéseket célzó köz- és magánforrások központosított gyűjtésére és megfelelő hatékonyságú és igazságosságú elosztására;
6. Egy globális MI-adat keretrendszer kialakítása, amelynek célja az újabb rendszerek és modellek tréningezését szolgáló adattömegek felhasználásával kapcsolatos szabályozási szttenderdek, jó gyakorlatok megosztása;
7. Egy különálló MI Hivatal létesítése az ENSZ Főtitkár Hivatalán belül, az MI-vel kapcsolatos kérdések legmagasabb szintű képviselőjére, illetve a sokrétű témakör átfogó szemléletű kezelésére.

[The UN unveils plan for AI](#)

## Első ízben engedélyezték, hogy egy MI-chatbot befektetési tanácsokat adjon

Mindeddig példa nélkül álló és a szakemberek által súlyos kockázatokat felvető módon immár hivatalosan, legálisan is szolgáltathat piaci befektetési tanácsokat egy mesterséges intelligencia. Az izraeli pénzügyi szabályozó hatóság engedélye alapján egy MI-technológiákra szakosodott startup – az ország egyik nagy bankjával összefogva – tőzsdeműveletekkel kapcsolatos piaci tanácsadói szolgáltatás indít – algoritmusokra alapozva a döntéseket. Az ellenállás nem véletlen, és ennek megfelelően a pénzügyi világ egyes szakértői nyomban megkondították a vészharangot, mondván, hogy amennyiben a mesterséges intelligencia technológiák használata széles körben válik elterjedtté a tőzsdék világában, úgy az rövid időn belül drámaian fogja szétzilálni a piacok működését. A pénz- és tőkepiacok hatósági szabályozásért felelős Israel Securities Authority közelmúltbeli döntése feljogosítja a Tel-Avivi székhelyű Bridgewise vállalatot arra, hogy még szeptember hónapban élesben is bevezesse Bridget névre keresztelt MI-chatbotját. Az előfizetéses alapon működő chatbot tőzsdei részvények eladásával és vásárlásával kapcsolatos befektetési tanácsokat fog nyújtani a felhasználóknak. Az új szolgáltatás bevezetésére a helyi startup az országban kiterjedt hálózattal rendelkező Israel Discount Bankkal szövetkezett. Miközben az OpenAI cég ChatGPT modelljének nyilvános debütálása óta eltelt időben a világ pénzügyi színtereitől sokan kezdtek bele a technológia széleskörű használatába, ez alapvetően a piackutatás, illetve az ügyfélkapcsolatok területére szorítkozott. Szakértők ugyanis nagyon komoly aggályokat fogalmaztak meg azzal kapcsolatban, hogy, amennyiben nagyon sok piaci szereplő fog ugyanazon algoritmusok ugyanazon adatkészleteire támaszkodva részvényügyi döntéseket hozni, az óhatatlanul a piac összeomlásához, súlyos pénzügyi válsághoz vezethet majd.

[A Chatbot is Now Cleared to Give, Buy, Sell Investment Advice](#)

## Hogyan győzzünk a generatív MI chatbotokkal szemben?

A Financial Time egy érdekes cikkének szerzője, Aswath Damodaran a pénzügyek professzora, aki szakmája miatt gyakran beszél, ír vállalati pénzügyekről. Egy barátja készített egy chatbotot, amely feldolgozta a szerző minden írását, minden vállalatértékelést, ami valaha is közzétett. Most az történik, hogy az így megalkotott beszélgetőrobottal elkezdtek vállalatokat – pénzügyi szempontból – értékelni, és ugyanezt a feladatot kapták a professzor diákjai is. Maga Damodaran sem tudja, hogy milyen eredményt szeretne látni: ki/mi lesz a jobb a feladatban. Addig is, amíg az eredmény ismert nem lesz, – a pénzügyi szakmára koncentrálna – megosztotta véleményét az MI-rendszerek lehetséges szerepéről és jövőjéről. Az MI területén kezdőként úgy látja, hogy három helyzetben lehet előnye a gépi rendszernek: a mechanikus, nem pedig az intuitív munkában; a szabályalapú, nem pedig az elvi alapú tudományágakban; és az olyan feladatokban, ahol a szubjektív ítéletek helyett objektív válasz létezik. Mit tehetünk annak érdekében, hogy megnehezítsük a munka gépek általi kiszervezését vagy az MI-vel való helyettesítését? Milyen lesz az a világ, amelyben az MI-re való támaszkodás általános lesz? A szerző négy gondolatot oszt meg ezekkel a kérdésekkel kapcsolatban. 1. A gépi rendszerek elterjedése segíteni fogja azokat a szakembereket, akik generalisták, azaz több szakterületen is otthonosan mozognak, és képesek átlátni a nagy egészet. 2. A befektetés és a vállalatértékelés terén, ha az értékelés technikája elsősorban pénzügyi modellezésen alapszik, akkor a múltbeli adatok

extrapolálásával, az MI gyorsabban és sokkal kevesebb hibával tudja az emberi munkát elvégezni. Viszont ha az értékeléshez „puha”, számszerűleg nem kifejezhető információ (üzleti történet) is szükséges, akkor ez nehézséget fog okozni az MI-rendszer számára. 3. El kell kerülni, hogy olyan kérdések megválaszolására használjuk az algoritmust, amelyekre magunknak kellene kitalálnunk a választ. Ha magunk csináljuk, az hosszabb ideig tarthat, és rossz válaszokhoz vezethet, de ez egy megtanult készség, amelyet elveszíthetünk, ha hagyjuk elszorvadni. 4. Az elme pillanatnyi üressége pedig lehetővé teszi azt, hogy látszólag összefüggéstelen tények összekapcsolódjanak, és az „Aha” pillanatok megélésének képessége csak az emberekre jellemző. A lényeg, hogy az emberi tevékenység mennyit tesz hozzá az adott feladat elvégzéséhez. Ez dönti el, hogy mennyire lesz védve a munkánk a gépi rendszerek térhódításától.

### [Strategies to beat the AI bots](#)

### **Hagyjuk csak, hogy a sztochasztikus papagájok kedvükre rikácsoljanak a szaksajtóban**

A „sztochasztikus papagáj” kifejezést a nagy nyelvi modellek kritikai jellemzésére használják. Az LLM-ek jelenléte észrevehető immár a tudományos szakfolyóiratokhoz közlésre benyújtott publikációkban is. A szakmai reakció eddig túlnyomórészt negatív volt, azaz elítélték és betiltották, hogy a modellek társszerzőként jelenjenek meg. Az „AI and Ethics” szakfolyóirat egyik friss írásának szerzője azonban, rendkívül részletes szakirodalmi anyagot feldolgozó cikkében megkérdőjelezi ezt a gyakorlatot. Azt állítja, hogy a folyóiratoknak lehetővé kellene tenniük, hogy az LLM-ek a cikkek társszerzőiként tűnjenek fel, feltéve, hogy jelentős mértékben használták ezeket a rendszereket az adott cikk megírásához. „Jelentős mértékű használaton” a szerző a következőket érti: számottevő mértékű hozzájárulás a mű koncepciójához vagy megtervezéséhez, vagy a műhöz szükséges adatok beszerzéséhez, elemzéséhez vagy értelmezéséhez. Ide tartozik továbbá az is, ha jelentős mértékben használták a gépi modellt a munka szerkesztéséhez, illetve a fontos intellektuális tartalom kritikai felülvizsgálatához. E feltételek teljesülése esetén engedélyezett lehetne a nagy nyelvi modellek használata, sőt társszerzőként való feltüntetése is. Ennek azonban lennének bizonyos kritériumai. A humán társszerzőknek tényellenőrzést kell végezniük az összes LLM által generált tényállításon, ennek megkönnyítése érdekében a humán társszerzőknek az LLM-től független irodalmi áttekintést kellene végezniük. A humán társszerzőket erősen ösztönözni kell arra is, hogy tekintélyes forrásokat találjanak minden LLM által generált ötlethez, amely nem tényállítás. A humán társszerzőknek továbbá minden LLM számítást és elemzést kétszeresen ellenőrizniük kell. A humán társszerzőknek vállalniuk kell azt is, hogy teljes mértékben elszámoltathatóak legyenek az LLM minden hozzájárulásáért. A humán társszerzőknek közzé kell tenniük, hogy melyik LLM-et használták, mikor használták, és hogy ez a használat hogyan járult hozzá a műhöz. Végül pedig a humán társszerzőknek minden szerzői joga és jogi felelőssége megmaradna a fentebb vázolt módon készült cikkekénél.

### [Let stochastic parrots squawk: why academic journals should allow large language models coauthor articles](#)





## Trendek

### Személyiség-bizonylatok ember mivoltunk online igazolására

Nemrégiben még mémek visszatérő fordulata volt, hogy az online terekben, webhelyeken és platformokon való mozgásunk során immár nem csak valamilyen captcha-séma segítségével kell majd bizonyítanunk, hogy mi bizony nem vagyunk robotok. A generatív MI-technológia szédületesen gyors terjedésével egyre inkább élő problémává válik ember voltunk bizonyításának módja, technikája. Ahogyan az MI-rendszerek egyre tökéletesebben képesek utánozni az embert, egyre nehezebbé válik különbséget tenni a tényleges emberi tevékenységek online formái, illetve az MI által teremtett mimikák között. A kérdés pedig fontos, sőt életbevágóan nagy jelentőségű lehet számtalan esetben, helyzetben. Bár juthatna eszünkbe fontosabb – személyiségi jogokhoz, egzisztenciához, vagyonbiztonsághoz, de akár élethez fűződő – eshetőség, amikor lényeges annak eldöntése, hogy valós ember megnyilvánulásival állunk-e szemben, a MIT Technology Review cikkében most a nagypolitika egyik sok aggodalmat kiváltó témáját említik: a dezinformáció kérdéskörében a szerzőség gépi/emberi volta sorsdöntő lehet. Most tudósok egy csoportja (amely az OpenAI, a Microsoft, a MIT és a Harvard 32 kutatójából áll) egy olyan koncepciót dolgozott ki, amely alkalmas lehet a probléma orvoslására. Az általuk javasolt személyazonosság hitelesítő rendszer („personhood credentials”) – amely csupán a legszükségesebb személyes adatokat vonná be a verifikálásba – arra épít, hogy az MI-rendszerek ma még nem képesek kijátszani bizonyos kriprográfiai eljárásokat, illetve, hogy az ember személyes, offline megerősítő akciói révén csalhatatlan módon bizonyítható egy-egy online szereplő gépi-, vagy emberi mivolta.

[How Personhood Credentials could help prove you're a human online](#)

## A mesterséges intelligenciával segített plagizálás problematikája: trendek és megoldások

A Nature folyóirat cikke a „gépi plágium áradat”, vagyis a nagy nyelvi modellek segítségével előállított szövegek (és persze képek vagy hanganyagok) szerzőségével, eredetiségével, szellemi tulajdonlásával kapcsolatos kérdéshalmazzal foglalkozik. A szerző, Diana Kwon igyekszik szisztematikusan, alaposan, és a megannyi leágazást is számba véve felrajzolni az LLM-támogatta plagizálás rendkívül sokrétű problematikáját. Az írás jó áttekintés az elsősorban a felsőoktatási-tudományos közvéleményt erősen foglalkoztató, kiterjedt és nagy horderejű következményekkel járó jelenségről. A cikk sok forrásból összegyűjtött, számszerűsített formában mutatja be a jelenség fejlődését és ezek a beszédes adatok egyértelműen az ilyen gyakorlatok növekvő tendenciáját mutatják. A szerző a cikkében erősen hangsúlyozza a kérdés nehezen eldönthető mivoltát: a különféle MI-használatok a tudományos munka során nagyon nehezen kategorizálhatók egy egyértelmű etikai doboz-rendszerbe. Ennek megfelelően óriási a bizonytalanság az ilyen „tudományos munka elkészítése során való MI-használat” kezelésével, szabályozásával kapcsolatban és hogy miért nincsen könnyű, egységes válasz a követendő politikák kérdésében. A témakör szinte minden fontos aspektusát figyelembe vevő cikk fontos erénye az is, hogy igyekszik alaposan bemutatni az „ellenszer” keresés jelenlegi irányait és nehézségeit egyaránt. A ChatGPT nyilvános bemutatkozását követően szinte az első időktől kezdve jelentkeznek a piacon a – hasonlóan az MI-technológia potenciáljait csatasorba állító – „MI-plágium detektorok”. A Nature írása hangsúlyosan fogalmaz: ezek az eszközök nem csupán technikai értelemben megbízhatatlanok és így vállalt funkciójukra alkalmatlanok. Néhány detektálási jellegzetességük folytán komoly tudománypolitikai feszültségeket is teremthet alkalmazásuk, például mert tendenciózusan „csalásnak” minősítik a nem angol anyanyelvűek által írt szövegek tekintélyes részét.

[AI is complicating plagiarism. How should scientists respond?](#)

## A generatív MI-rendszerek elszennyezik az emberi tudást megtestesítő adatokat

2024. szeptember elején jelentette be a wordfreq program fejlesztője, hogy a továbbiakban nem fogja a szoftvert frissíteni. A wordfreq egy Python könyvtár a szavak gyakoriságának keresésére számos nyelvben, számos adatforrás alapján. A rendszer segítségével elemezni lehet a változó nyelvi szokásokat, lényegében tehát a nyelvelőződés folyamatát. Az OSCAR (Open Super-large Crawled Aggregated coRpus) nyílt forráskódú projekten alapszik, amelynek célja, hogy webalapú, többnyelvű adathalmazokat (Wikipédia, cikkek, könyvek, weboldalak, Twitter és Reddit, azaz több millió forrást) biztosítson a gépi tanulás (ML) és MI-alkalmazások számára. A megálmodói kifejezetten azt szeretnék elérni, hogy nagy mennyiségű kommentálatlan, nyers adat álljon rendelkezésre. Azaz az emberi nyelvnek egy olyan lenyomata, amely azt tükrözi, hogy az emberek hogyan használják a nyelvet. A wordfreq vezetője szerint azonban a 2021-es év utáni gyűjtés, már olyan mértékben torzult a nagy nyelvi modellekből származó egyre nagyobb tömegű szöveges információ miatt, hogy nem lehet többé tiszta képet kapni a tényleges nyelvallapotot illetően. Ahogy a fejlesztő megfogalmazta: „Nem hiszem, hogy bárkinek is megbízható információi lennének a 2021 utáni emberi nyelvhasználatról. A nyílt web (az OSCAR-on keresztül) volt a wordfreq egyik adatforrása. Most az egész web tele van nagy nyelvi modellek által generált, senki által nem írt, semmit nem kommunikáló szlogenekkel. Ha ezt a moslékot beleszámítjuk az adatokba, az elferdíti a szófrekvenciákat.” – fogalmaz keserűen a projektgazda. Persze, a wordfreq adatforrásaiban is előfordult spam, de kezelhető és gyakran azonosítható módon.



A nagy nyelvi modellek azonban olyan szöveget generálnak, amely valódi nyelvnek álcázza magát és a kimenetük mindenhol felbukkan.

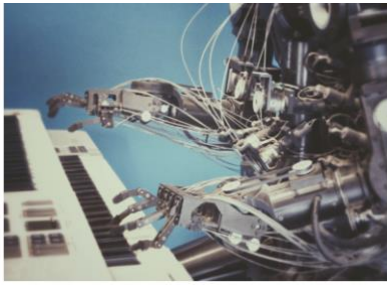
[Why wordfreq will not be updated](#)

### **Az MI-rendszerekben testet öltő innováció lenyűgöző, de még nem emelte a termelékenységet**

Régóta folyik az éles vita arról, hogy a mesterséges intelligenciának milyen hatása van, illetve lesz a foglalkoztatottságra, a termelékenységre. Most az amerikai szövetségi bankrendszer vezetői fejtették ki nézeteiket egy konferencián. A vélemények hátterét az az ellentét adja, amely az utóbbi 15 év kifejezetten szerény amerikai termelékenység növekedése és az MI-rendszerek fejlődésének legutóbbi szakaszát kísérő felfokozott kockázati tőkebefektetési tevékenység között feszül. A bankárok elismerik a mesterséges intelligenciában rejlő hosszú távú lehetőségeket a gazdasági növekedés ösztönzésére, de ez egyelőre csak ígélet. Az MI-rendszerek látványos fejlődése, a fejlesztésekre és a kiszolgáló infrastruktúrára fordított tekintélyes összegek ellenére a termelékenység nem nagyon emelkedett. Érzékelhető ugyan némi inkrementális emelkedés, és ebben már szerepet játszanak a különböző típusú MI-rendszerek is. De az még külön vizsgálat tárgya, hogy specifikusan a generatív MI-rendszerek hozzájárulása a termelékenység növekedéséhez hasonlóan mérsékelt lesz-e, vagy látványosabb eredményekkel kecsegtet majd. A számok nyelvén megfogalmazva: az amerikai termelékenység 2,7%-kal nőtt a második negyedévben a tavalyi év azonos időszakához képest, de egyébként a század eleje óta évente átlagosan körülbelül 1,5%-kal emelkedett, dacára az informatika meglehetősen látványos fejlődésének. Azt még túl korai volna kijelenteni, hogy hosszútávon fenntartható, növekvő termelékenységet jelentenek ezek a friss számok. Abban viszont nincs vita, hogy az MI és különösen a generatív MI valószínűleg általános célú technológiává válik, amely elterjed az egész gazdaságban. Ezek a fejlesztések megteremthetik a nagyobb termelékenység, és így a stabil árakkal összhangban lévő magasabb bérnövekedés feltételeit. Egyelőre azonban sok vállalat csak „kóstolgatja” az MI-rendszereket. Kísérleteznek az alkalmazásával, de tartózkodnak attól, hogy nagyobb mértékben befektessenek a technológiába. Így az MI még nem befolyásolja jelentősen a foglalkoztatást és a bevételeket.

[Fed's Cook says 'impressive' AI innovation has yet to lift productivity](#)





## Működésben

### „A ChatGPT miatt hagytam ott a tanítást”

A népszerű Time magazinban megjelent, drámai című esszé szerzője szépíró, irodalomtudós, pedagógus. Pontosabban csak volt pedagógus. Legutóbbi munkahelyén, egy műszaki felsőoktatási intézményben doktori hallgatóknak tanította a tudományos kutatóktól megkövetelt írás fortélyait. Azóta abbahagyta a tanítást, mert hibaválónak bizonyult az a törekvése, hogy a hallgatói értelmesen, képességeiket és készségeiket ténylegesen építő módon használják a nagy nyelvi modelleket (ChatGPT). A diákok, műszaki emberként tökéletesen tisztában voltak a ChatGPT hiányosságaival. A szerző pedig nem naív elutasítással közelítette meg a problémát. Mégis minden erőfeszítése ellenére hallgatói teljesen mechanikusan használták, a lehető legrövidebb idő és a legkisebb erőfeszítés ráfordításával oldották meg a feladatokat, az eredménytől függetlenül. A szerző problémája nem az, hogy a ChatGPT által létrehozott szöveget nem lehet felismerni. A problémát abban látja, hogy a ChatGPT-vel segített írást könnyű azonosítani, de nehéz bizonyítani. Ennek eredményeképpen sok órát töltött olyan írások osztályozásával, hiányosságai javításával, amelyekről tudta, hogy ChatGPT-vel készültek. Úgy látja, hogy több időt töltött azzal, hogy visszajelzést ad az MI-rendszernek, mint a diákjainak. A problémák egyik része abból adódott, hogy az MI-rendszer hatékonyan javította a nyelvtani hibákat (különösen a külföldi diákok esetében), de eközben finom módon megváltoztatta a szöveget, és a hallgatók nyelvi műveltsége kevés volt ennek felismeréséhez. Ahogy egyikük megfogalmazta, a ChatGPT által javított szöveg hatásosabb, magával ragadóbb. A másik probléma az, hogy az írás nem pusztán az író fejében lévő tartalmak áttevése egy másik médiumba, hanem a gondolkodással szorosan összefüggő folyamat. Nem egyszerű és könnyű munka. A szerző tapasztalata szerint sok - talán a legtöbb - diákja már nem hajlandó ezzel a „kellemetlenséggel” megküzdeni, inkább a könnyebb utat választja, és így elveszítik a lehetőséget, hogy mélyebben átgondolják a kutatásaikat.

### [I Quit Teaching Because of ChatGPT](#)

## A nyelvi modellek saját valóságérzékelést fejlesztenek ki

A világ egyik legnevesebb felsőoktatási és kutató intézménye, a Massachusetts Institute of Technology egy nemrégiben közzétett cikkében egy közelmúltban lezárult, igen figyelemre méltó kutatásról számolt be. A szigorú normák szerint tervezett és kivitelezett, ellenőrzött kísérletként futó kutatásban arra derült fény, hogy mélyen az LLM-eken belül kifejlődhetnek olyan „valóság szimulációk”, amelyek túlmutatnak az egyszerű mimikri viselkedésen. Ezek a jelenségek pedig a nyelv mélyebb megértéséről tanúskodnak. Ez a felfedezés kisebb sokkot okoz, különösen az olyan nem-technikai szakembereknek, akik mindaddig abban a hitben élnek, hogy a mára már napi tevékenységeinkben szerepet kapó MI, a nagy nyelvi modell valójában egy nyelvtisztítási alapon működő szoftver, amely semmit sem „ért”, csupán az óriási képzési adattömegre támaszkodva, a soron következő írásjelek valószínűségét elemzi, illetve becüli meg. A MIT Számítástechnikai és Mesterséges Intelligencia Laboratóriumának (CSAIL) munkatársai ezzel szemben arra jutottak, hogy a nagy nyelvi modellek képesek lehetnek saját valóságérzékelést (understanding of reality) kifejleszteni annak érdekében, hogy ezzel javítsák a generatív képességeiket. Tehát, valami olyasmit, amit eddig egyértelműen és alapvetően az emberi képesség attribútumaként könyveltünk el. „Ez a kutatás közvetlenül a modern mesterséges intelligencia egyik központi kérdését célozza meg: a nagy nyelvi modellek meglepő képességei egyszerűen a méretarányos statisztikai összefüggéseknek köszönhetőek, vagy a nagy nyelvi modellek értelmes megértést alakítanak ki arról a valóságról, amellyel dolgozniuk kell? Ez a kutatás azt jelzi, hogy az LLM belső modellt fejleszt a szimulált valóságról, még akkor is, ha soha nem képezték ki arra, hogy ezt a modellt fejlessze” - mondja Martin Rinard, az MIT EECS professzora, a CSAIL tagja és a tanulmány vezető szerzője.

[LLMs develop their own understanding of reality as their language abilities improve](#)

## Forradalmi áttörést hozhat az infarktus korai előrejelzésében egy MI

Egy Nagy-Britanniában most kidolgozott vizsgálati eljárás forradalmi áttörést ígér a fejlett világban szinte mindenütt népbetegségnek számító infarktus, illetve a hasonló „21. századi járványként” elkönnyelhető cukorbetegség, vagy stroke veszélyeinek korai előrejelzésében. A brit állami egészségügyi szolgálat, az NHS (National Health Services) által finanszírozott fejlesztésben az Oxfordi Egyetem spinoff cége, a Caristo Diagnostics dolgoz ki olyan eljárást, amely mesterséges intelligencia technológiák segítségével elemzi a – diagnosztikában persze már hosszú ideje ismert és alkalmazott – CT-felvételeket. Az újítás lényege abban áll, hogy a súlyos elváltozásokhoz vezető biológiai folyamatok korai jeleit képes azonosítani a rendszer, így tulajdonképpen a CT-felvételek szokásos vizuális kiértékelésével nem azonosítható veszélytényezők időben történő leleplezésére alkalmas. A jelenleg öt angliai kórházra kiterjedő pilot projektben a startup cég által kifejlesztett CaRi-Heart MI-platform az infarktusos állapotra utaló, kezdeti tünetekkel érkező páciensek CT-felvételeit vizsgálva azonosít olyan gyulladáshoz vezető kórképeket, amelyek a veszélyes betegség későbbi megjelenésére utalnak. Az algoritmus által felállított diagnózist természetesen képzett orvosok ellenőrzik, kiszűrve az esetleges tévedéseket. A probléma horderejét jól mutatja, hogy Nagy-Britanniában mintegy 7,6 millió ember élhet valamilyen szívbetegséggel, és évente nagyjából 350 000 betegről készítenek CT-vizsgálatot, infarktusos tünetekkel, illetve gyanúval. A hagyományos elemzési módszerek mellett jelentős arányban „eredménytelen” vizsgálatok összköltsége pedig mintegy 7,4 milliárd fontsterling költséggel terheli az egészségpénztárt.

[‘Game changer’ AI detects hidden heart attack risk](#)

## A Google MI-szoftvere fogja segíteni a munkanélküli juttatások elbírálását Arizonában

Nevada állam a következő hónapokban egy olyan, a Google által működtetett, generatív mesterséges intelligencia-rendszer bevezetését tervezi, amely megoldaná a munkanélküliségi segély megítélésével kapcsolatban felgyülemlett nagy számú fellebbezés problémáját. Nagyon sok jegyzőkönyv feldolgozásáról van szó, amelyek a fellebbviteli meghallgatások során keletkeztek. A maga nemében ez lesz az első eset az országban, hogy egy ilyen rendszer jelentős hatással bíró közigazgatási döntést befolyásolhasson. A várakozások szerint a Google rendszere felgyorsítja a fellebbezési eljárást: a határozat megírásához szükséges idő több órától egyes esetekben mindössze öt percre csökkenhet. A rendszer a meghallgatási jegyzőkönyvek és a bizonyító erejű dokumentumok alapján ajánlásokat készít, és saját elemzéssel látja el az adott személy munkanélküliségi kérelmének jóváhagyását, elutasítását vagy módosítását. Ezt követően legalább egy emberi bíró minden egyes ajánlást felülvizsgál. Ha a bíró egyetért az ajánlással, akkor aláírja és kiadja a határozatot. Ha a bíró nem ért egyet, akkor felülvizsgálja a dokumentumot, és az illetékes állami munkaügyi szervezet (Department of Employment, Training & Rehabilitation – DETR) kivizsgálja az eltérést. Christopher Sewell, a DETR igazgatója hangsúlyozza, hogy nem lesz olyan a gépi rendszer által létre hozott írásbeli döntés, amely emberi beavatkozás és emberi felülvizsgálat nélkül lép hatályba. Számosan szakértő azonban aggodalmát fejezte ki, hogy az ügyintézés sebességére helyezett hangsúly alááshatja az emberi tényező korlátozó, de a jogbiztonságot megtestesítő szerepét. Technikailag a rendszer úgy fog működni, hogy egy általánosan elérhető alapmodellt finomhangolnak (retrieval-augmented generation, azaz RAG-technológiával) a nevadai munkaügyi szabályozást és a korábbi fellebbviteli ügyeket tartalmazó adatbázis alapján.

[Google's AI will help decide whether unemployed workers get benefits](#)

