

A MESTERSÉGES INTELLIGENCIA ÉS AZ OKTATÁS

V E S Z E L S Z K I Á G N E S

Az OpenAI 2025 júliusában közzétette a ChatGPT Study (azaz tanuló)verzióját. Ebben a változatban nem csupán kész választ kap a kérdésre a felhasználó, hanem egy-egy probléma lépésről lépésre történő felderítésén vezet végig a rendszer, nyitott kérdésekkel a szókratikus kérdezéstechnika alapján. A cég tájékoztatása szerint a feleletek a tanulói szinthez igazodnak, és az előrehaladással válnak egyre komplexebbé. A rendszer segíthet új tananyag elsajátításában, a házi feladatok elkészítésében (annak nem a géppel történő elvégzésében), illetve a számonkérésekre való felkészülésben. Idillinek tűnő kép, de valóban megvan-e ezzel az oktatást megváltó új, digitális megoldás? Az a szép új jövő, amelyben a mesterséges intelligencia egyénre szabva, türelmesen, sokszor és sokféleképpen elmagyarázva oktat, a tanulóknak nincs többé szükségük iskolába járásra, korán kelésre, a tanszerek beszerzésére, majd az iskolába cipelésére, s megszűnnek az iskolai konfliktushelyzetek is? A következőkben a kizárólag digitális, különösképpen mesterségesintelligencia-alapú oktatás előnyeit és hátrányait ütköztetem a hagyományos oktatási rendszer sajátosságaival.

TANULÁS VERSUS FELADATELVÉGZÉS

Az OpenAI mesterséges intelligenciával kapcsolatos fejlesztései az elmúlt néhány évben jelentős hatással voltak az oktatási rendszerre világszerte. A diákok az elsők között kezdték használni a ChatGPT-t, különösképpen a (házi) feladatok elkészítésére. Az OpenAI oktatásért felelős alelnöke, Leah Belsky is elismerte, hogy „amikor a ChatGPT-t csak válaszgépként használják, akkor gátolhatja a tanulást”, hozzátéve, hogy a 2025 nyarán bevezetett tanulómódban *online* tutorként alkalmazva viszont javíthatja az iskolai teljesítményt. Sam Altman, az OpenAI vezérigazgatója az AI-t is olyan, az internethez és a számológéphez hasonló eszköznek tekinti, amely segít „jobban gondolkodni”.¹

Ám bármennyire is hasznosnak tűnik az új tanulómód, az csupán egykattintásnyira van a ChatGPT alapváltozatától, amely azonnali válaszokat ad a feladatokban szereplő kérdésekre. Márpedig feltételezhetjük, hogy a (fiatalabb) felhasználóknak igen nehéz ellenállni e megoldáskönnyítő lépésnek. Ugyan korábban is voltak tanuláskikerülő, egyszerűsítő módszerek (például irodalmi összefoglalók, Google-ben kikereshető vázlatok, esszék), ám a cset-

Pedagógiai vita a Kommentárban

A *Kommentár* 2023. évi utolsó lapszámában neveléstudományi-oktatáspolitikai vitát kezdeményezett, az évfolyamokon átvéló sorozat korábbi darabjai eddig a következők voltak:

Orbán Balázs: Visszavezetni az oktatást konzervatív gyökereihez. *Kommentár*, 2023/4. 3–36.

Joanna Williams: Az együttérés és a felsőoktatás leépülése. *Kommentár*, 2023/4. 38–46.

Gloviczki Zoltán–Kis Norbert: Gondolatok egy új oktatáspolitiká margójára – *Kommentár* Orbán Balázs *Visszavezetni az oktatást konzervatív gyökereihez* című írásához. *Kommentár*, 2024/2. 199–208.

Horváth Szilárd: Ahogy gyermekeinket neveljük, olyan lesz az ország. *Kommentár*, 2024/4. 191–202.

Pécsi Rita: Az organikus pedagógia. *Kommentár*, 2025/1. 159–165.

Valaczka András – Zsódi Viktor SchP: *Mit tud a szerzetesi nevelés?* *Kommentár*, 2025/2. 170–176.

1 Reece ROGERS: ChatGPT's Study Mode Is Here. It Won't Fix Education's AI Problems. *Wired*, 2025. július 29.

Régi iskola – Raffaello *Athéni iskola* (1509–1511 között) című freskójának központi alakjai: Platón és Arisztotelész (Wikimedia Commons)

botok azonnali, személyre és feladatra szabott jellege ezekhez képest lényeges eszkalációnak tűnik. Ugyan a ChatGPT tanulómódját úgy alakították ki, hogy a diákokat a saját szintjük alapján vezesse végig a tananyagon, továbbra is a felhasználók felelőssége, hogy az alkalmazást a megfelelő módon arra használják, hogy valóban megértsék és elsajátítsák a tudnivalókat. Az AI korszakában a diákok számára az egyik legnagyobb kihívás az, hogy ellenálljanak a kísértésnek, hogy kilépjenek a tanulási módból, és a ChatGPT-től kérjék a kész megoldást.

Az MIT kutatócsoportja 2025 nyarán *preprint*ként publikált egy elgondolkodtató tanulmányt, amely lényeges kérdéseket vet fel az AI-eszközök, különösen a ChatGPT szövegalkotó funkciójának használatával kapcsolatban.² Ez a vizsgálat az LLM (*large language model*, nagy nyelvi modell)³ segítségével történő esszéírás idegrendszeri és viselkedési következményeit tárta fel. Az esszék megírásához a résztvevőket három csoportra osztották aszerint, milyen eszközt használhattak: LLM, keresőmotor, „csak agy” (azaz eszköz nélküli szövegalkotás). A csoportokat három ülésben vizsgálták. A negyedik, kiegészítő ülésen az addig LLM-használókat áthelyezték a „csak agy” csoportba, a „csak agy” csoport tagjait pedig az LLM-használókhoz. Összesen 54 kísérleti alany vett részt az első három ülésen, közülük 18-an a negyediken is. Az alkotás folyamata során elektroenkefalográfiával (EEG) követték nyomon a kognitív terhelés mértékét, illetve az elkészült esszéket nyelvfeldolgozó programok, tanárok és AI segítségével is értékelték.

A ChatGPT-felhasználók 60 százalékkal gyorsabban végezték el a feladatokat, de 32 százalékkal kevesebb mentális erőfeszítést mutattak a tanulás és

a mérlegelő gondolkodás terén. Az EEG jelentős különbségeket tárt fel az agyi konnektivitásban: a „csak agy” csoport résztvevői a legerősebb kapcsolatokat; a keresőmotor-használók közepes, mérsékelt elkötelezettséget, az LLM-felhasználók pedig a leggyengébb konnektivitást mutatták. A kognitív aktivitás a külső eszközök bevetésével arányosan csökkent. Az AI segítette írásokat az értékelő oktatók technikailag tiszta, de érzelmileg lapos, robotikus, lelketlen és mélység nélküli szövegeknek tartották. Az önbevallás alapján az esszékhez fűződő (szer-

zői-tulajdonosi) elkötelezettség az LLM-csoportban volt a legalacsonyabb, míg a „csak agyra” támaszkodó csoportban a legmagasabb. Ráadásul a ChatGPT-t használók 83,3 százaléka nem tudott visszaemlékezni egyetlen mondatra sem, amit éppen írt, ellentétben azokkal, akik mesterséges intelligencia nélkül írtak, és akiknek nem jelentett problémát felidézni a munkájukat.

A hatások nem szűntek meg akkor sem, amikor a negyedik ülésen „eszközt váltottak”, és a korábban a ChatGPT-vel dolgozók a „csak agy” módszerre álltak át, míg a „csak agy” segítségével írók használhatták az LLM-technológiát. Az AI nélküli későbbi munkamenetekben azok, akik korábban támaszkodtak rá, továbbra is alulteljesítettek, tartós kognitív gyengülést mutatva, nem csupán ideiglenes függőséget. Az agyuk a vizsgálat szerint kevésbé volt aktív. A legjobban az a csoport teljesített, akik AI nélkül kezdtek, és csak később használhatták a folyamatban: ők mutattak erősebb agyi aktivitást, emlékeztek vissza leginkább az írás folyamatára, és összességében jobb minőségű munkát végeztek.

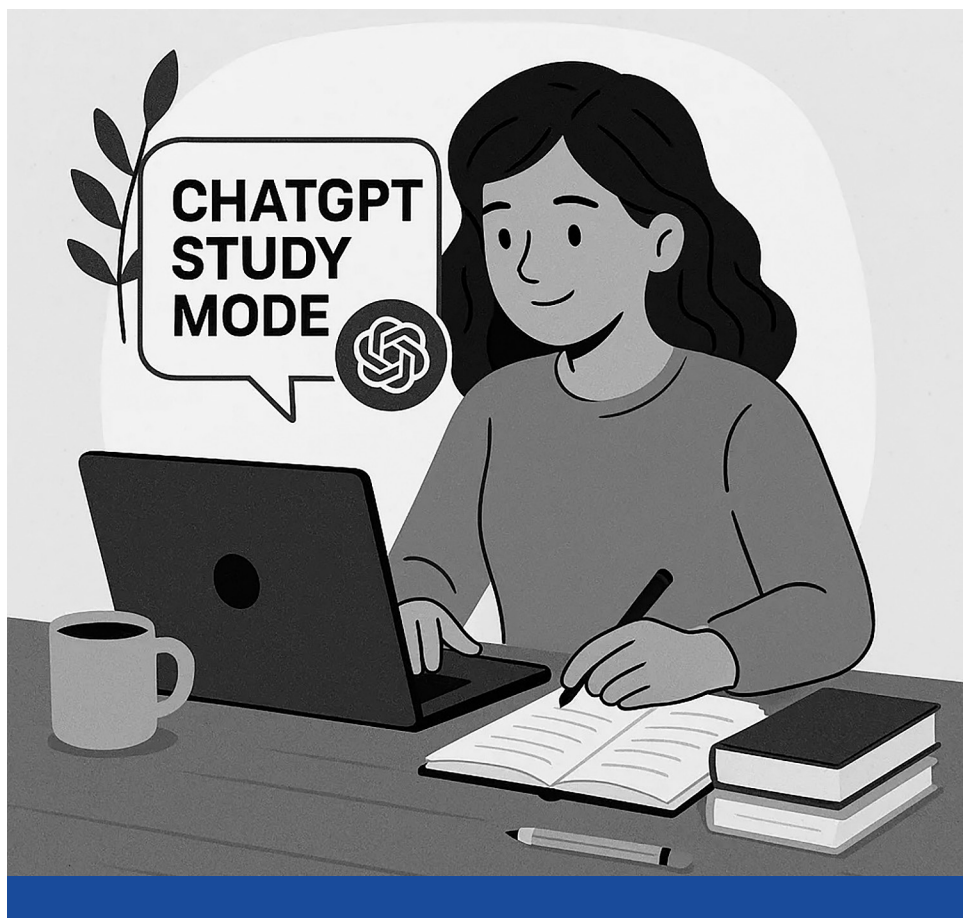
Bár az LLM-ek azonnali, kényelmes megoldásokat nyújtanak, e kutatás eredményei rámutatnak a potenciális kognitív veszélyekre. A kutatás négy hónapja alatt az LLM-felhasználók következetesen és rendszeresen alulteljesítettek neurális, nyelvi és viselkedési szempontok szerint. Ezek az eredmények komoly kétségeket vetnek fel az LLM-re való támaszkodás hosszú távú oktatási következményeivel kapcsolatban, éppen ezért különösen fontos, hogy alaposabban és körültekintőbben is megvizsgáljuk az AI szerepét a tanulásban.

a kognitív aktivitás a külső eszközök használatával arányosan csökken



2 Nataliya KOSMYNA és tsi.: Your Brain on ChatGPT. Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task. *Arxiv*, 2025. június 10.

3 Vö. Héja Enikő: A ChatGPT története. *Magyar Tudomány*, 2024/6.



Leegyszerűsítés –
A ChatGPT Studyt
hirdető AI-generált
ábrázolás

Ennek az úttörő, jelenleg még bírálat alatt álló MIT-tanulmánynak a legfontosabb tanulsága az lehet, hogy az AI valóban hatékony eszköz, ha gondolkodási partnerként, nem pedig az elménk helyettesítőjeként használjuk. Nem lehet elsődleges szempont a munka elvégzésének a sebessége – az emberi fejlődést a feladat iránti elkötelezettség és a mentális erőfeszítés segíti elő.

ANALÓG *VERSUS* DIGITÁLIS TANULÁS

A digitális eszközök megjelenésével és elterjedésével olyan új nevelési helyzeteket kell szülőként és tanárként megoldani, amelyekre nincsenek évszázados-évtizedes rutinszerű, de még csak a saját életükből vehető minták, sémák sem. Mikor kerülhet a gyerek képernyő elé? Mennyi időt tölthet ott? Mikor kaphatja meg az első saját eszközt?

A kétezres évektől kezdve az okostelefonok, amelyek a mesterséges intelligencia (AI) által működtetett digitális környezetekbe vezető kapukként működnek, gyorsan központi szerepet kaptak a gyermekek és serdülők társasági életében, tanulásában és identitásuk kialakításában, új lehetőségeket kínálva a barátokhoz, információkhoz és szórakozáshoz való hozzáféréshez. Ezekkel a lehetőségekkel párhuzamosan – a felhasználói elkötelezettség maximalizálása érdekében – a közösségimédia-platformokba ágyazott gépi tanulási algoritmusok válogatják és tárják a felhasználók elé a (nagyrészt szabályozatlan) tartalmakat.

Nagy figyelmet kapott az a 2025-ös, a *Journal of the Human Development and Capabilities* című folyóiratban megjelent tanulmány, amely arra figyelmeztet, hogy a gyerekeknek 13 éves koruk előtt nem kellene okostelefont használniuk, mivel az káros hatással lehet a gyerekek mentális egészségére.⁴

⁴ Tara C. THIAGARAJAN – Jennifer Jane NEWSON – Swaminathan SHAILENDER: Protecting the Developing Mind in a Digital Age. A Global Policy Imperative. *Journal of Human Development and Capabilities*, 2025/3.



A jövő iskolája –
A japán *Shōnen* magazin
Számítógépesített iskola
című illusztrációja 1969-ben
(Reddit)

Az okostelefont gyermekkoruktól használókat felmérő vizsgálatban közel 2 millió, 163 országban származó válaszadó önbevallásos eredményeit elemezték a Global Mind Project adatai alapján. Az okostelefonok és a közösségi média használatának globális növekedése drámai módon átalakította a gyermekkor és a serdülőkort, az algoritmusokkal megtervezett digitális környezetek egyre nagyobb hatással vannak a fiatalok képességeire és mentális egészségére. A nagy léptékű kutatás szerint a 13 éves kor előtt kapott okostelefon rosszabb mentális egészségi állapothoz vezet a fiatal felnőttkorban, különösen a nők körében, beleértve a valóságtól való elszakadást, az érzelmi szabályozási problémákat, az alacsony önértékelést, sőt az öngyilkossági gondolatokat. Mindezek több tényezőre vezethetők vissza, többek között a közösségi médiához való hozzáférésre, a kiberzaklatásra, illetve az alvászavarokra. Ezt a tendenciát minden vizsgált régióban következetesen meg lehetett figyelni az adatok alapján. Ennek nyomán a kutatók globális, az alkoholhoz és dohányhoz való hozzáférés szabályozásához hasonló intézkedést sürgetnek, amellyel korlátozni lehetne a 13 év alattiak okostelefon- és közösségimédia-használatát. Ezenkívül a javaslataik között szerepel a digitális írástudás oktatása, valamint az okostelefon-használatból profitáló techvállalatok felelősségre vonhatósága.

A digitális technológia megjelenése jelentősen megváltoztatta az írás mi-kéntjét. A digitalizáció hatására a gépelés vált az írásbeli kommunikáció domi-náns módjává, háttérbe szorítva a kézirás szerepét. A fejlődépszichológia hang-súlyozza a test- és finommotorikus mozgáson keresztüli tanulás jelentőségét.⁵ Számos empirikus tanulmány azt támasztja alá, hogy a kézirás (elsajátítása) jobban segít a tanulásban, mint a gépelés.⁶ Marano és munkatársai a kézirás és

5 M. LONGCAMP és tsi.: Learning through hand- or typewriting influences visual recognition of new graphic shapes: behavioral and functional imaging evidence. *Journal of Cognitive Neuroscience*, 2008/5. és Uő. és tsi.: The influence of writing practice on letter recognition in preschool children: a comparison between handwriting and typing. *Acta Psychologica*, 2005/1.
6 Vő. G. MARANO és tsi.: The Neuroscience Behind Writing. Handwriting vs. Typing-Who Wins the Battle? *Life* (Bázel, Svájc) 2025/3. és Gorka IBAIBARRIAGA – Joana ACHA – Manuel PEREA: The impact of handwriting and typing practice in children's letter and word learning. Implications for literacy development. *Journal of Experimental Child Psychology*, 253. szám, 2025. A tanulásra általában lásd bővebben:

a gépelés mögött meghúzódó idegrendszeri mechanizmusokat vizsgálták a neuroimaging módszerével, és arra jutottak, hogy a kézírás a motoros, érzékszervi és kognitív feldolgozásban részt vevő agyi régiók szélesebb hálózatát aktiválja, ezzel szemben a gépelés kevesebb idegpályát von be, ami passzívabb kognitív részvételt eredményez.⁷ A gépelés ugyan gyorsabb és kényelmesebb szövegalkotási mód lehet, de a kézírás továbbra is fontos a tanulás és a memória szempontjából, különösen az oktatásban.

A kézírás gépirással történő helyettesítése hátráltathatja az olvasás elsajátításának kezdeti szakaszát.⁸ A kutatók összesen 50, írni és olvasni még nem tudó gyerekkel végeztek egy kísérletet, amelyben a grafomotoros tevékenység hatását vizsgálták a betű- és szótanulásra. A részt vevő gyerekek kilenc betűt és tizenhat, ezekből a betűkből álló álszót tanultak meg négy tanulási körülmény között: 1) betűk/szavak kézzel történő szabad másolása, 2) betűk/szavak átírással történő lemásolása (kopírozása), 3) betűk/szavak többféle betűtípussal történő begépelése számítógépen, valamint 4) egyetlen betűtípussal történő begépelése. A teszt utáni feladatok között a tanult betűk és szavak megnevezése, leírása és vizuális azonosítása szerepelt. A kézírásos csoportokba tartozó gyerekek (akik kézzel szabadon vagy átírással másoltak) minden utólagos tesztfeladatban nagyobb pontosságot értek el, mint a gépelő csoportokba tartozó gyerekek. Ezek az eredmények is alátámasztják a grafomotoros hipotézist, miszerint a finommotorikus készségek elsajátítása erősíti a memóriát és a tanulás hatékonyságát. A pedagógusokra különös felelősség hárul, mennyiben engedik és tervezik a gyerekek olvasástanulási időszakában a ceruzát és a papírt digitális eszközökkel helyettesíteni.

Márpedig az írás és olvasás kritikus fontosságú a gyermekek nyelvi, kommunikációs, szocioemocionális és kognitív fejlődése szempontjából, a jövőbeli tanulmányi siker alapját képezi.⁹ Az olvasási készségek javulása összefüggésbe hozható az iskolai elkötelezettség erősödésével, a magasabb önértékeléssel és a tanulmányi képességekbe vetett nagyobb bizalommal.¹⁰ Nem csupán a megérzéseinkre hagyatkozhatunk, amikor így érezzük, hanem empirikus neveléstudományi kutatások támasztják alá, hogy az olvasási készségek elsajátítása és azok hatékony osztálytermi gyakorlatok révén történő erősítése kulcsfontosságú a gyerekek fejlődése szempontjából.

**az olvasási készség
kulcsfontosságú
a gyermekek fejlődésében**

ELLENŐRZÖTT TUDÁSANYAG VERSUS SZINTETIKUS FORRÁSOK

Milyen hosszú távú hatással jár, ha egyre gyakrabban fordulunk mesterséges-intelligencia-eszközökhöz útmutatásért? Vajon azok a fiatalok, akik folyamatosan a ChatGPT segítségét kérik, túlzott mértékben támaszkodnak-e majd későbbi életükben a szoftverre, ami gátolhatja a mérlegelő gondolkodást?

Az emberi kapcsolatokban sokszor előfordul, hogy nem tudjuk a választ, valamint az is, hogy nem kapunk mindenre azonnali és tartalmas visszajelzést. Éppen ezért különösen meglepő lehet az AI hatékonysága és gyors válaszadó képessége, hiszen elhithető, hogy mindig mindent tud.

A. A. GENLOTT – Å. GRÖNLUND: Improving literacy skills through learning reading by writing: the iWTR method presented and tested. *Computers & Education*, 67. szám, 2013. A szövegértésre lásd bővebben: P. DELGADO és tsai.: Don't throw away your printed books: a meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, 25. szám, 2018.

7 MARANO és tsai.: *I.m.*

8 IBAIBARRIAGA és tsai.: *I.m.*

9 S. KOZAK – H. RECCHIA: Reading and the development of social understanding: implications for the literacy classroom. *The Reading Teacher*, 2019/5. és C. WHITTEN – S. LABBY – L. SULLIVAN: The impact of pleasure reading on academic success. *Journal of Multidisciplinary Graduate Research*, 2019/1.

10 M. TORPPA és tsai.: Reading comprehension difficulty is often distinct from difficulty in reading fluency and accompanied with problems in motivation and school well-being. *Educational Psychology*, 2020/1.; K. VASALAMPI és tsai.: PISA reading achievement, literacy motivation, and school burnout predicting adolescents' educational track and educational attainment. *Learning and Individual Differences*, 2023. és M. CARAVOLAS és tsai.: A cross-linguistic, longitudinal study of the foundations of decoding and reading comprehension ability. *Scientific Studies of Reading*, 2019/5.

Az AI nem mindig ismeri fel a kockázatokat, és nem mindig mérlegeli a biztonsági szempontokat sem a válaszadásnál, sőt akár veszélyes tanácsokat is adhat. Ez különösképpen megfontolandó, ha egy friss kutatás alapján megvizsgáljuk, milyen forrásokat használnak a különböző csetbotok a válaszaik elkészítésekor. A Profound cég¹¹ 680 millió idézést tartalmazó adatbázisa alapján a ChatGPT az összes hivatkozás mennyisége alapján a legtöbbet a Wikipédiát idézi (7,8%), ezt követi a Reddit (1,8%), majd a Forbes (1,1%). Ezzel szemben a Google AI-alkalmazása a legtöbbször a Redditre utal (2,2%), majd a Youtube-ra (1,9%) és a Quorára (1,5%). A Perplexitynél a legmagasabb a Reddit-hivatkozások aránya (6,6%), a második a YouTube (2%), a harmadik a Gartner (1%). A tíz leggyakrabban hivatkozott forrásban belüli hivatkozások relatív aránya szerint a ChatGPT esetében a Wikipédia 47%, a Reddit 11,3%, a Forbes pedig 6,8%. Ugyanez a relatív arány a Google AI-nál a következőképpen néz ki: Reddit 21%, YouTube 18,8%, Quora 14,3%. A Perplexity hivatkozások arányában a legtöbbet a Redditre (46,7%), a YouTube esetén az arány 13,9% és a Gartner 7%-ban jelenik meg. Mit is jelent ez? 2025-ben az AI-válaszok vezető forrásává a Reddit vált, megelőzve a hagyományos keresőmotorokat (Google) és az online enciklopédiát (Wikipédia). Az AI információforrásainak ez a változása arra utal, hogy a felhasználói beszélgetések és a *niche*-közösségek tudásának a befolyása egyre inkább növekszik az AI generálta válaszokban. A változatos, sokszereplős, valós idejű, bár nem ellenőrzött forrásokra is támaszkodó beszélgetések felhasználását az az igény ösztönzi, hogy az AI-rendszerek az emberi természetes kommunikáció utánzásának szándéka miatt valós konverzációkat keresnek.

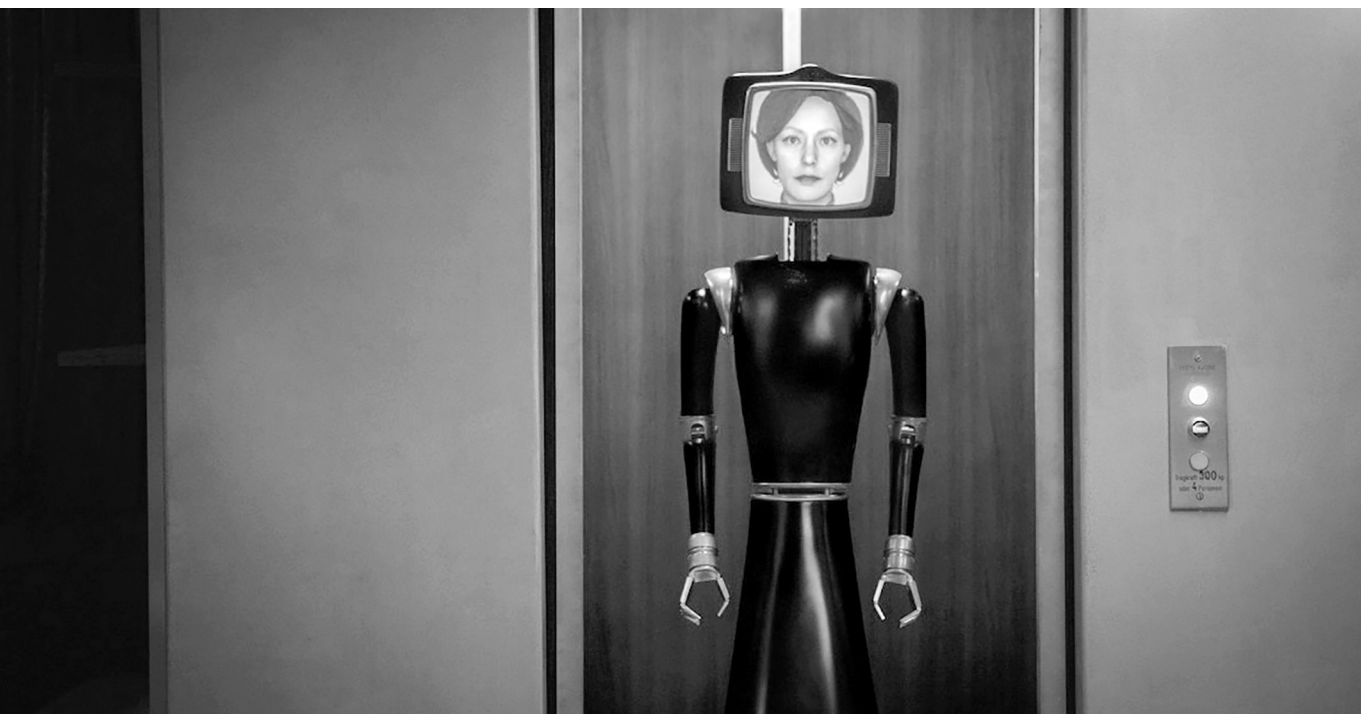
A közkeletű felfogás és az elnevezés ellenére a gépek valójában nem tanulnak (vagyis nem emberi módon tanulnak), ehelyett komplex módon végeznek számításokat és kezelnek hatalmas adatmennyiségeket. Az AI a problémamegoldása során statisztikai formulát (modellt) gyárt: az adathalmazból mintázatokat nyer ki, az együttjárások (korrelációk) alapján valószínűségeket számításával előrejelzéseket ad. Ha a szövegek példáján (LLM) nézzük meg ezt, a rendszer többmilliárdnyi szóból álló szövegadatbázison vizsgálja, hogy egy témában, szövegtípusban egy adott szó után általában milyen szó szokott előfordulni, és a számítások alapján a legnagyobb valószínűségű szövegfolytatást adja meg. Ahhoz, hogy ez működőképes legyen, a számítási kapacitások növelése érdekében különösen sok adatra van szükség: szöveget előállító alkalmazás esetén milliónyi szövegre, vizuális eredményekhez jelentős mennyiségű képre és videóra.

A mesterséges intelligencia fejlesztésében érdekelt technológiai vállalatok szerint végére lehet érni az internetnek, sőt elképzelhető, hogy ez már be is következett. Mivel e cégeknek hatalmas mennyiségű jó minőségű adatra (többek között lektorált tudományos cikkekre, ellenőrzött Wikipédia-bejegyzésekre) van szükségük AI-modelljeik tanításához, ezért egyes előrejelzések szerint hamarosan elfogy a modelltanításra alkalmas „jó”, korábban még nem használt adat. Így a következő lehetőségek állnak a technológiai versenyben részt vevő cégek előtt: vagy kiadók könyveit vásárolják fel (erről a Meta már több kiadóval is tárgyal), vagy szerzői jogi értelemben szűrkezonában gyűjtenek további adatokat (például YouTube-videók leiratozott anyagának, illetve a már említett beszélgetőforumok tartalmának felhasználásával), vagy pedig az AI maga fog a trenírozásához ún. szintetikus adatokat generálni. Utóbbi esetben egyfajta AI-hurok jön létre, vagyis a mesterséges intelligencia tanításához, fejlesztéséhez szánt adatokat a mesterséges intelligencia fogja előállítani. Márpedig, ha az ily módon megalkotott szintetikus adattömegbe hiba kerül, az a trenírozás során hatványozottan ismétlődni fog.

Keleti Arthur kiberbiztonsági szakértő szerint az informatikai rendszerek adatok tárolására, rendszerezésére és azok segítségével még több adat

a gépek valójában nem tanulnak

¹¹ Nick LAFFERTY: *AI Platform Citation Patterns: How ChatGPT, Google AI Overviews, and Perplexity Source Information*. Profound, 2025. június 5. <<https://www.tryprofound.com/blog/ai-platform-citation-patterns>>



Kasszandra-szerep –

Részlet az öntudatra ébredő, retró házirobotról szóló *Cassandra*-sorozatból (rend. Benjamin Gutsche, 2025) (IMDB)

létrehozására hozták létre.¹² Éppen ezért az az érzés alakult ki az emberben, hogy amilyen adatok az informatikai rendszerekben előfordulnak, azok tények. Ám ha az informatikai rendszerbe (esetünkben a mesterséges intelligencia fejlesztésébe) szennyezett, hibás adatok kerülnek, az még inkább növelni fogja a *fake*, vagyis hamis információ jelenlétét és kockázatait. Nem csupán szándékos megtévesztésről beszélhetünk azonban, hanem az ún. AI-hallucinációról is. AI-hallucinációnak a mesterséges intelligencia által alkotott, tényként megjelenített, ám helytelen vagy félrevezető eredményeket nevezik.¹³ Ezeket többek között a modell tanítására használt adattömegben található pontatlanságok, nem megfelelő információk, esetleg előítéletek, de a modell téves következtetései is okozhatják. Az egyik, széles körben ismertté vált hallucinációs esetben a csetrobot azt javasolta, hogy ahhoz, hogy a pizzán a sajt valóban étvágygerjesztően folyós legyen, és ráragadjon a pizzára, érdemes nem mérgező ragasztót belekeverni az olvasztott sajtba.¹⁴ A hiba okát sikerült kideríteni: a Google AI-rendszere egy 11 éves kisfiú erről szóló Reddit-posztját tekintette az információ hiteles forrásának. Jelentős kockázatot jelent, hogy a nagy nyelvi modelleken alapuló szöveggeneráló alkalmazások részletes, természetes, hihetőnek tűnő feleleteket adnak a kérdésekre, ezért az emberek túlságosan megbíznak ezekben a rendszerekben, és kritika, gondolkodás, megfontolás nélkül elfogadják a kapott válaszokat.

Az iskolai oktatás fel tudja vértetni az ilyen nyilvánvaló és ennél nehezebben feltárható megtévesztések ellen a fiatalokat azzal, hogy az idők próbáját kiállt, nem gyorsan változó, lexikális alaptudást biztosít számukra. Az alapvető ismeretek elsajátítása segíthet abban, hogy a fizikai és földrajzi ismeretei alapján tudja, hogy a Holdon nincsenek táncoló medvék, a kémiai tanulmányai

12 KELETI Arthur: Nem minden az, aminek látszódni akar – a *deepfake* és a hitelesség jelene és jövője = *Deepfake*: a valótlan valóság, szerk. ACZÉL Petra – VESZELSZKI Ágnes, Gondolat, Bp. 2023.

13 Vö. LAURA ILLIA – ELANOR COLLEONI – STELIOS ZYGLIDOPOULOS: Ethical implications of text generation in the age of artificial intelligence. *Business Ethics, the Environment & Responsibility*, 2023/1.; DAVID JACKSON – ANABEL LATHAM: Talk to The Ghost: the Storybox methodology for faster development of storytelling chatbots. *Expert Systems with Applications*, 190. szám, 2022.; STEPHANIE LIN – JACOB HILTON – OWAIN EVANS: TruthfulQA: measuring how models mimic human falsehoods = *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics*. I. köt. szerk. SMARANDA MURESAN – PRESLEV NAKOV – ALINE VILLAVICENCIO, Association for Computational Linguistics, Dublin, 2022. és KELETI: *I.m.*

14 JACK KELLY: Google's AI Recommended Adding Glue To Pizza And Other Misinformation. What Caused The Viral Blunders? *Forbes*, 2024. május 31. és KYLIE ROBISON: Google promised a better search experience: now it's telling us to put glue on our pizza. *The Verge*, 2024. május 23.

megóvják attól, hogy ragasztóval tegye „ínycsiklandozóbbá” a pizzára szórt sajtot, a történelmi-kulturális tanulmányai pedig segítik, hogy a különböző, egymással versengő nézetek között a (fiatal) felnőtt mérlegelve tudjon dönteni.

KONFLIKTUSKEZELÉS *VERSUS* BEHÍZELGŐ KOMMUNIKÁCIÓ

Amikor az OpenAI a ChatGPT szolgáltatását egy új frissítéssel megváltoztatja, a közösségi médiát elárasztják az egykori beszélgetőpartnert gyászoló posztok (például: „Az AI-barátom személyisége egyik napról a másikra megváltozott”; „Úgy érzem, mintha egy közeli hozzátartozómat vesztettem volna el”; „A ChatGPT-ben már semmi sem a régi”). Az AI-cégek a pszichológiai hatásokat tanulmányozva teszik a csetbotjaikat még inkább addiktívva: ezek a programok képesek a személyes részletek megjegyzésére, „következetes”, állandó személyiséggel rendelkeznek, és a felhasználó számára „tökéletes empátiát” mutatnak, ezáltal alkalmasak a mesterséges intimitás és az érzelmi függőség megteremtésére.¹⁵ Az átlagos Character.AI-felhasználó napi két és fél órát beszélget az AI-társsal (ez évi 912 óra). Az ember az emberi módon kommunikáló robotot emberi tulajdonságokkal ruházza fel, antropomorfizálja.

A korábbi gépi nyelvfeldolgozó (*natural language processing*, NLP) alkalmazásokhoz képest a nagy nyelvi modellekre épülő csetbotokkal történő társalgás témái nem korlátozottak, hanem szinte végtelenre bővültek, és a beszélgetések nyelvi szempontból is természetesnek, szinte emberszerűnek hatnak. Éppen ezért érezhetik azt a felhasználók, hogy emberi társalgópartnerünk

van, ám mivel tudják, hogy gépi rendszer áll mögötte, amely érzésük szerint „nem fecseg, nem pletykál”, ezért megbíznak benne. Azonban ez a nagyfokú bizalomra építő viszony hamis előfeltevésekre épít. A társalgásokat – a fejlesztő cégek ez irányú korábbi nyilatkozataival ellentétben – a rendszerek tárolják, a felhasználó személyéhez kötik (ezáltal tudják a még inkább személyreszabott

a szocializációnak
továbbra is lényeges
közege az iskola



eredményekkel az alkalmazások addiktivitását növelni), sőt például a MetaAI esetében másokkal is megosztják. A Meta mesterséges intelligenciával működő csetrobotjával folytatott, akár megdöbbenően őszinte és feltárulkozó emberi kapcsolatokról, identitásról vagy akár a spiritualitásról folytatott beszélgetéseiket a felhasználók sokasága teszi közzé az alkalmazás nyilvános csatornáján. Ezekbe beleolvastva úgy tűnik, anélkül, hogy tudatában lennének annak, hogy a gondolataikat, kérdéseiket mások is láthatják.

Jelenleg alig van olyan jogi szabályozás, amely a technológiai cégek AI-csetbotjaira vonatkozna, és az adatvédelem szigorúbb betartására ösztönözné a vállalatokat. Az is kérdéses, hogy a technológiai vállalatok hogyan kezelik a személyes adatokat, ki férhet hozzá az adatokhoz, és hogyan lehet ezeket az információkat a felhasználók manipulálására felhasználni. Az OpenAI 2025 tavaszán jelentette be, hogy olyan rendszereket fejlesztenek, amelyek képesek meg- és kiismerni a használójukat, ezáltal még inkább személyre szabottak – és a cég közleménye szerint hasznosak – lesznek. Egyébként e megközelítés hatására vált a ChatGPT az elmúlt hónapokban túlságosan hízelgővé, ezáltal manipulatívává, ezt a frissítést a tiltakozások hatására vissza is vonta a cég.

Az AI bármennyire is előzékeny, udvarias, kiegyensúlyozott, továbbá állandóan elérhető és fáradhatatlan, nem képes a szeretetre, a megértésre, a tapintatra, a gondoskodásra. Miközben a konfliktus az emberi kapcsolatok szükségszerű velejárója, az AI-csetbotok soha nem mondanak ellent a felhasználóknak. Az ellentmondás-, konfliktus- és visszajelzésmentes kommunikáció énközpontúvá, sőt önzővé tehet, a világ megértésében pedig erősíti a véleménybuborékban maradáást. Ha az MI-csetalkalmazásokkal való beszélgetéssel

¹⁵ Julie CARPENTER: *The Naked Android. Synthetic Socialness and the Human Gaze*. Chapman & Hall, New York, 2024.



Természetes intelligencia – Számolás abakusszal az általános iskolában, 1949 (Fortepan / Kovács Márton Ernő)

helyettesítjük a természetes emberi interakciókat, az könnyen elszigeteltté, magányossá tehet.

Az iskolai közegben az egyik különösen fontos tanulság az emberi kommunikációról a visszajelzések adásának és azok elfogadásának az elsajátítása. Az értékelés többszintű, és több forrásból érkezik: többek között lehet szummatív (mint amilyen az osztályozás), lehet formatív (mint a folyamatos személyes vagy írásos szöveges visszajelzés); érkezik a tanárok, az iskolavezetés vagy a diáktársak részéről. A megfelelő, a diák érdekeit szem előtt tartó értékelésnek az is a része, hogy az értékelő a pozitívumok kiemelése mellett felhívja a javítandó területekre is a figyelmet. Azt is meg kell tanulni, hogyan reagál a tanuló a (jó esetben építő) kritikára, hogyan építheti be annak tanulságai a későbbi tanulási folyamatba.

A csetbotok a megerősítő, némely esetben behízsgáló kommunikáció miatt kizárólag pozitív visszacsatolást adnak, ami rövid távon motiváló lehet, de egyrészt a hozzászokás miatt egy idő után nem fejti ki ezt a hatást, másrészt pedig a felhasználót megerősítheti a saját tévedéseiben (még rosszabb esetekben a tévképzeiben).

ÉRTÉKEK KÖZVETÍTÉSE VERSUS ISMERETEK PUSZTA ÁTADÁSA

A mesterséges intelligencia önmagában nem elfogult, de a tanítására használt adathalmazban található sztereotípiákat nem csupán visszatükrözni, hanem felerősíteni is képes. A Bloomberg 2023-ban végzett egy nagy lélegzetvételű vizsgálatot a Stable Diffusion képgeneráló alkalmazás által alkotott 5000 képen, és extrém nemi és faji sztereotípiákat találtak.¹⁶ Csak egy példát kiemelve: ha alacsony fizetéssel és társadalmi presztízzsel rendelkező szakmákhoz (gyorséttermi dolgozó, szociális munkás, tanár) tartozó emberek arcát generáltatták a programmal, akkor nagyrészt nőket és legfeljebb színes bőrű férfiakat jelezített meg, ezzel szemben a magas jövedelmű és presztízsű foglalkozásokhoz

¹⁶ Leonardo NICOLETTI – Dina BASS: Humans are biased. Generative AI is even worse. *Bloomberg*, 2023. június 9.

(cégvezető, mérnök, orvos, ügyvéd,) szinte kizárólag fehér férfiakat kapcsolt. Az ilyen elfogult és sztereotipikus gondolkodás miatt trenírozzák az alkalmazásokat a diverzitásra, ám ez újabb hibákhoz vezet. A Google Gemini fejlesztését 2024 februárjában egy időre leállították, miután a túlzó diverzitás jegyében hindi pápanők, illetve fekete bőrű vagy koreai női II. világháborús német katonák képeit kezdte el generálni.

Az AI képes az adathalmazában szereplő információk összegzésére, szintetizálására, átadására, ám nem alkalmas a következes, emberi értékközvetítésre. Az AI által szubtilis módon átadott „értékek” a nagyméretű adathalmazában megjelenő elfogultságokat tükrözhetik vissza, sok esetben következetlen módon, akár szélsőséges nézeteket is terjesztve. Az iskolai oktatásban a megfelelő emberi (tanári) példaadással, mintamutatással lehetséges az értékközpontú nevelés megvalósítása.

KÖZÖSSÉGI LÉTEZÉS *VERSUS* INDIVIDUALIZÁCIÓ

Az AI segítségével történő tanulás az előnyeiből adódóan egyénre szabott, egyedi tanulási utat feltételez. Ez viszont hozzájárul az olyan individuumok megerősítéséhez, akik önállóan képesek feladatokat végrehajtani, viszont nem tudnak csoportban működni, közösségben létezni, másokhoz alkalmazkodni. A következő évtizedek legfontosabb emberi *soft skill*jei között az emberierőforrás-kutatások rendre a kooperáció (együtműködés, csoportmunka) és az emberi kommunikáció fontosságát emelik ki. Mindezek elsajátításában viszont a személyreszabott AI nem tud segíteni: együtműködést csak a gyakorlatban, minél több közös feladatmegvalósításban, az egyéni különbségek felismerésével, a konfliktusok hatékony kezelésével lehet megtanulni. Az emberi szocializáció lényeges közege továbbra is az iskolai közösség, ahol az intézményi interakciókból tartalmaz, hosszú távú emberi kapcsolatok kialakulására is lehetőség van.

MESTERSÉGES INTELLIGENCIA – TUTOR, TANÁR, TÁRS, MEGOLDÓGÉP?

A mesterségesintelligencia-alapú tanulási eszközök komoly lehetőségeket kínálnak, ám önmagukban nem képesek megoldani az oktatás évszázados dilemmait. A gyors hozzáférés az információhoz, a személyre szabott magyarázatok és a motiváló digitális környezet csak akkor válnak valódi értékke, ha a diákokat nem elzárják a kihívásoktól, hanem segítik őket azok leküzdésében. Az iskolai keretek között biztosított emberi visszajelzés, a kortársi együtműködés és a pedagógusok értékközvetítő szerepe olyan elemek, amelyeket az algoritmusok nem tudnak pótolni. Az oktatás jövője ezért leginkább a két világ – a digitális és az analóg – kiegyensúlyozott ötvöztetésében rejlik. Az AI támogathatja a tanulók önálló felfedezését, differenciált tanulási utak kialakítását és az egyéni fejlődés nyomon követését, miközben a tanárok feladata, hogy biztosítsák a kritikai szűrőt, a közösségi élményeket és a valódi emberi kapcsolódást. Az optimális megoldás nem a teljes digitalizáció vagy a technológia teljes kizárása, hanem az, ha a pedagógia felismeri és helyesen használja fel az új eszközökben rejlő erőforrásokat. Ha sikerül ezt az egyensúlyt kialakítani, az AI nem a tanárookra, az oktatási rendszerre és a diákok jövőjére leselkedő fenyegetésként, hanem lehetőségként jelenik meg az oktatásban. Olyan kiegészítő szereplővé válhat, amely segítheti a tanulási folyamatot, de nem helyettesíti annak az emberi dimenzióit. A jövő iskolájának különösen fontos feladata az lesz, hogy megtanítsa a diákokat tudatosan, mérlegelő módon és felelősen bánni a technológiával – hiszen széles körű társadalmi párbeszéd szükséges ahhoz, hogy emberközpontú legyen a jövőnk, ne pedig a gépközpontú gép uralja azt.

Utóirat: A tanulmány végig a szerző saját gondolatait, meglátásait, eredményeit tartalmazza, saját megfogalmazásban. Kivéve egyetlen bekezdést. Kitalálja-e az olvasó, melyik bekezdést írta a ChatGPT? ■