

A GENERATIV MESTERSÉGES INTELLIGENCIA, MINT TUDÁSTRANSZFER ESZKÖZ: TANULSÁGOK MARKETING MSC VIZSGAFELADATOK BENCHMARKOLÁSÁBÓL

Tétényi Bálint – Kovács Stefán – Csapó Anna – Kádár-Buksa Márk

Összefoglalás

A kutatás célja annak vizsgálata volt, hogy az egyetemek és munkahelyek között miért alakul ki időbeli késés a mesterséges intelligencia (MI) és digitális technológiák átvételében, valamint hogyan lehet ezt mérsékelni. Eredményeink szerint a vállalatok gyorsabban alkalmazzák az új eszközöket, míg az oktatási intézményekben az akkreditációs és szervezeti korlátok lassítják az adaptációt. A hallgatók gyakran önállóan vagy munkahelyi környezetben sajátítják el a szükséges digitális készségeket, míg az oktatók informális csatornákból értesülnek az új trendekről. A tudástranszfer gyorsítása érdekében kulcsfontosságú a felsőoktatás és a munkaadók közötti együttműködés, közös kurzusok, mentorprogramok és szakmai partnerségek kialakítása. A tanulmány rámutat, hogy a digitális eszközök integrációja hozzájárulhat a pályakezdeők felkészítéséhez a munkaerőpiac aktuális igényeire.

Kulcsszavak: Mesterséges Intelligencia, generatív MI, tudástranszfer, marketing, LLM, digitális technológiák, készségfejlesztés

JEL: D83

GENERATIVE ARTIFICIAL INTELLIGENCE AS KNOWLEDGE TRANSFER TOOL: LESSONS FROM BENCHMARKING MARKETING MSC EXAM TASKS

Abstract

The study examined why a time lag occurs between universities and workplaces in adopting artificial intelligence (AI) and digital technologies, and how this gap can be reduced. Our findings show that companies adopt new tools more rapidly, while educational institutions face accreditation and organizational barriers that slow adaptation. Students often acquire essential digital skills independently or in workplace settings, while lecturers frequently rely on informal channels to learn about emerging trends. To accelerate knowledge transfer, stronger collaboration between higher education and employers is essential, including joint course development, mentoring schemes, and professional partnerships. The research highlights that integrating AI and digital tools into academic curricula can better align graduates' skills with labor market demands, particularly supporting early-career professionals.

Keywords: Artificial Intelligence, generative AI, knowledge transfer, marketing, LLM, digital technologies, skills development

Bevezetés

A generatív mesterséges intelligencia (MI) térnyerése alapvetően formálja át a szervezeti tudásátadás folyamatait: lehetővé teszi nagy mennyiségű explicit tudás gyors összegyűjtését, összefoglalását és a felhasználó igényeihez igazított közvetítését. A szakirodalom szerint a generatív modellek új tartalmakat állítanak elő a rendelkezésre álló adatok mintázatainak tanulása alapján (Pedersen, 2023), a chatbotok pedig szöveges vagy hangalapú interakcióban képesek támogatni a tudástranszfert és a beilleszkedési folyamatokat (Adamopoulou és Moussiades, 2020). A promptolás pontossága meghatározó tényező a releváns és megbízható kimenetek eléréséhez (Sivakumar et al., 2024). Emellett a generatív MI képes az információáramlás automatizálására, hosszú dokumentumok összefoglalására és személyre szabott tanulási útvonalak kialakítására (Toma et al., 2023), ami közvetlenül érinti a szervezeti tanulást és a tudáskoordinációt.

A szervezeti tudás szempontjából különbséget kell tennünk tacit és explicit tudás között (Chikán, 2008). Miközben a tacit tudás társas folyamatokon keresztül adódik át, a generatív MI különösen erős az explicit tudás externalizálásában és (újra)internalizálásában: képes a dokumentált ismeretekből összefüggő, feladatra szabott tudásegységeket képezni. Ugyanakkor a generált tartalmak azonosíthatósága és ellenőrizhetősége kulcsfontosságú a téves információk és a dezinformációk kockázatának kezelésében, ami a hiteles tudásátadás előfeltétele (Najee-Ullah et al., 2022). A személyre szabott támogatás erősítheti a felhasználói elköteleződést és motivációt is (Chan és Hu, 2023), de mindez csak megfelelő használati keretek mellett hoz tartós minőségi javulást.

Jelen tanulmány fókuszja: a generatív MI mint tudástranszfer-eszköz vizsgálata. Konkrétan arra keressük a választ, hogy különböző MI-platformok mennyire képesek meglévő, explicit tudást hatékonyan közvetíteni egy új felhasználó (pl. friss belépő munkatárs) számára. Ennek empirikus modellálására Budapesti Műszaki és Gazdaságtudományi Egyetem Marketing Msc szakon oktatott Marketing kutatás tantárgy anyagaira alapozva használtunk mint standardizált, komplex tudástartalmakat igénylő tesztkörnyezetet: a vizsgafeladatok alkalmasak arra, hogy a tudásátadás minőségét több feladattípuson (igaz-hamis, párosítás, esettanulmány, kifejtős) mérjük, és összevessük az MI teljesítményét különböző hozzáférési és tudásbetöltési feltételek mellett (pl. dokumentumok rendelkezésre állnak-e az MI számára vagy sem). Ezzel a keretkezéssel a hallgatói vizsga nem cél, hanem mérőeszköz: olyan próbatest, amelyen keresztül a szervezeti környezetben is releváns tudástranszfer-képességeket értékelhetjük.

A vizsgálat kapcsolódik, de nem azonos a pályakezdők onboarding folyamatának elemzésével. A jelen cikk elsődleges kérdése nem az, hogy a hallgatók vagy dolgozók hogyan „íráthatnak meg” feladatokat MI-vel, hanem az, hogy milyen feltételek mellett (pl. dokumentum-betöltés, hozzáférési szint, promptolási gyakorlat) képes az MI hitelesen és következetesen közvetíteni a szervezeti explicit tudást. A gyakorlati relevanciát jelzi, hogy ugyanezek a mechanizmusok alkalmazhatók a vállalati tudásmenedzsmentben is, például az alábbi – nem kizárólagos – felhasználásokban:

- belső szabályzatok, folyamatleírások és tudástárak kérdés-válasz alapú elérése és adaptív összefoglalása;
- esettanulmány-alapú helyzetértékelés és döntéstámogatás új munkatársak számára;
- tanulási útvonalak és mikro-tananyagok dinamikus előállítás meglévő dokumentumokból;
- minőségbiztosított tudásközvetítés (ellenőrzött forrásokból, hivatkozásokkal), a hibás vagy dezinformatív kimenetek kockázatának csökkentése.

A tanulmány hozzájárulása kettős. Egyrészt elméleti: a generatív MI-t a tudástranszfer irodalmi keretében értelmezi (explicit vs. tacit; externalizáció–internalizáció), és illeszti a promptolásról, személyre szabásról és információáramlásról szóló diskurzusba (Pedersen, 2023; Adamopoulou és Moussiades, 2020; Sivakumar et al., 2024; Toma et al., 2023; Chikán, 2008; Najee-Ullah et al., 2022; Chan és Hu, 2023). Másrészt empirikus: standardizált feladatstruktúráján keresztül vizsgálja, hogy a különböző MI-konfigurációk milyen mértékben képesek a meglévő tudás hatékony közvetítésére. A fókusz így következetesen a tudásátadás minőségén van; az onboarding, a friss diplomások betanítása és más HR/üzleti kontextusok lehetséges alkalmazási terepként jelennek meg, amelyekre az eredmények implikációkat fogalmaznak meg, de nem képezik a cikk közvetlen vizsgálati tárgyát.

Végül, a tanulmány tudatosan foglalkozik a megbízhatóság és hitelesség kérdéseivel. Míg a generatív MI skálázhatóvá teheti az explicit tudás elérését és személyre szabását, a tartalmi pontosság, a forrásellenőrzés és a dezinformáció kockázatának kezelése elengedhetetlen a szervezeti bevezetéshez (Najee-Ullah et al., 2022). Ennek megfelelően a további fejezetek a tudástranszfer szempontjából lényeges teljesítménymutatókra és a különféle MI-beállítások összehasonlítható hatásaira koncentrálnak.

Anyag és módszer

Hipotézisek

A kutatás központi kérdése az volt, hogy a generatív mesterséges intelligencia mennyire képes támogatni az explicit tudás közvetítését új felhasználók – például frissen belépő munkatársak – számára. A szervezeti tudástranszfer szimulációjához az MSc szintű vizsgafeladatokat tekintettük mérőeszköznek, mivel ezek komplex, többféle tudáselemet (lexikális ismeret, összefüggések felismerése, gyakorlati alkalmazás) igényelnek, hasonlóan az onboarding folyamatokhoz.

A vizsgálat során a következő hipotéziseket állítottuk fel:

- H1: Az ingyenes MI platformok kevesebb helyes választ adnak, mint a fizetős változataik, mivel korlátozottabb a kapacitásuk (pl. egyszerre csak korlátozott számú kérdésre képesek válaszolni).
- H2: Azok az MI rendszerek, amelyek dokumentumok betöltésére képesek, szignifikánsan jobb teljesítményt mutatnak, különösen az esettanulmányi típusú kérdések esetében, ahol az elméleti tudás gyakorlati alkalmazására van szükség.
- H3: Az általános tudásra épülő kifejtős kérdések esetében a legtöbb MI képes lesz helyes választ adni, de az igaz–hamis és a párosító feladatoknál jelentősek lesznek a teljesítménykülönbségek.

Kutatási design, vizsgálati minta

A vizsgálat alapját a Budapesti Műszaki és Gazdaságtudományi Egyetem Marketing MSc képzésének vizsgafeladatai képezték. Ezek a feladatok alkalmasak voltak a generatív MI-k tudásátadási képességeinek modellezésére, mivel többféle kérdéstípust tartalmaznak: igaz–hamis kérdések (lexikális tudás, tényellenőrzés), párosító feladatok (fogalmi kapcsolatok felismerése), esettanulmányi kérdések (komplex problémaelemzés, gyakorlati alkalmazás), kifejtős kérdések (szintetizáló tudás, önálló érvelés).

A kérdések forrásául egyetemi tananyagok szolgáltak, amelyek nem tartalmaztak bizalmas vállalati információt, ugyanakkor szimulálták a szervezeti tudás explicit formáit. Így a feladatok egyben modellként működtek arra, hogyan lehet az MI-t alkalmazni onboarding helyzetben: dokumentum alapú feldolgozás vs. saját tudásbázisra támaszkodás.

A kutatás során 6 MI platform teljesítményét mértük fel, amelyek eltértek hozzáférési feltételeikben (ingyenes/fizetős) és képességeikben (dokumentumfeltöltés engedélyezése vagy hiánya). Összesen 14 különböző konfiguráció került tesztelésre. A platformokat a 2024 áprilisában elérhető legfrissebb verziójukban vizsgáltuk.

A vizsgálat során minden platformnak 240 kérdésből álló teszt sorozatot adtunk, így összesen 3360 választ elemeztünk. A válaszokat az előre elkészített megoldókulcshoz viszonyítva értékeltük. Ha egy MI nem értette meg a feladatot, a kérdést újrafogalmaztuk, ami tükrözi az onboarding folyamat egyik kulcselemét: a kommunikáció pontosítása az új belépők számára.

A relevancia és helyesség mérésére kvantitatív elemzést alkalmaztunk. Az arányos helyességi mutatók mellett figyeltük a nyelvi minőséget (válasz magyar nyelven érkezett-e), valamint a megbízhatóságot (válasz konzisztenciája ismételt prompt esetén).

A „Towards Detection of AI-Generated Texts and Misinformation” (Faust et al., 2023) tanulmány alapján külön figyelmet fordítottunk a hibás vagy megtévesztő tartalmak azonosítására, mivel ezek a vállalati tudásátadásban kockázatot jelenthetnek. A kutatás nem törekszik végleges következtetések levonására, mivel az eszközök folyamatosan fejlődnek. Az MI platformok közötti különbségek elemzése kulcsfontosságú az MI technológiai fejlődésének megértéséhez (Aronsson et al., 2021).

1. táblázat: Kutatásban résztvevő MI platformok és jellemzőik

MI platformok	Fejlesztő	Alkalmazott nyelvi modell	Folyamatos Internet hozzáférés	Dokumentum feltöltési lehetőség	Havi előfizetési díj
ChatGPT-3.5	OpenAI	OpenAI GPT-3.5	nem	nem	0 USD
CHatGPT-4	OpenAI	OpenAI GPT-4	nem	igen	20 USD
Gemini	Google AI	LaMDA	igen	csak kép	0 USD
Gemini Advanced	Google AI	LaMDA	igen	csak kép	19.99 USD
Copilot	GitHub	OpenAI GPT-3.5	igen	igen	0 USD
Copilot Pro	GitHub	OpenAI GPT-4	igen	igen	10 USD
YouPro Genius	You.com	Open AI GPT-4	igen	igen	10 USD
Chatsonic	Writesonic	többféle	igen	pdf, word, kép, audio, url	12.67 USD
Perplexity	Perplexity	többféle	igen	igen	20 USD
ChatGPT-4 D.	OpenAI	Open AI GPT-4	nem	igen	20 USD
Chatsonic D.	Writesonic	többféle	igen	pdf, word, kép, audio, url	12.67 USD
Perplexity D.	Perplexity	Perplexity Language Model	igen	igen	20 USD
PDF AI PDF D.	Open AI	Open AI GPT-4	nem	igen	20 USD

Forrás: Saját kutatás alapján saját szerkesztés (2024)

A kutatásban 6 MI platform volt tesztelve, különbséget téve fizetős, nem fizetős használat, illetve dokumentum feltöltési lehetőséget biztosító verzió között, így összesen a kutatásban 14 generatív MI eszköz lett tesztelve. A kutatásban külső dokumentum feltöltését alkalmazó MI platformot D. megjelöléssel jelezzük (Például ChatGPT-4 dokumentum feltöltős verziója ChatGPT-4 D.).

Sokrétű alkalmazhatósága miatt az OpenAI által fejlesztett ChatGPT (<https://chatgpt.com>, megtekintve 2024. április 25.) termékek közül négy is bekerült a vizsgálatba, mivel van ingyenes hozzáférésű (ChatGPT-3.5), fizetős (ChatGPT-4) és dokumentum feltöltésére alkalmas (ChatGPT-4 D., PDF Ai PDF D.) eszköze is.

A Gemini (<https://gemini.google.com/app>, megtekintve 2024. április 25.), Gemini Advanced, illetve Copilot (<https://copilot.microsoft.com>, megtekintve 2024. április 25.) és Copilot Pro esetében két eszköz került be a kutatásba – az ingyenes (Gemini, Copilot) és a fizetős (Gemini Advanced, Copilot Pro) kategória.

A you.com (<https://you.com>, megtekintve 2024. április 25.) eszköz esetében két eszköz lett kiválasztva (YouPro Genius, YouPro Claude 3 Opus). Itt dokumentum feltöltési lehetőség van, azonban a kutatást nem sikerült lefuttatni a platformon, így a kutatásból a dokumentum feltöltős eleme ki lett zárva.

A Chatsonic (<https://app.writesonic.com>, megtekintve 2024. április 25.) és Perplexity.ai (<https://www.perplexity.ai>, megtekintve 2024. április 25.) két eszköze került bele a kutatásba (Chatsonic, Perplexity), melyek szintén szerepelnek a dokumentum feltöltésre alkalmas platformok között (Chatsonic D., Perplexity D.).

Kutatásból kizárt MI platformok listája az alábbi táblázatban megtalálhatóak (2. táblázat). A kizárások okai lehetnek a korlátozott kérdésszám/karakterszám (pdf.ai, Ingyenesen elérhető verziók: Chatsonic, Perplexity.ai, You.com Smart mód), nyelvi korlát (askyourpdf) és elérhetőségi korlát – Magyarországon nem elérhető MI platform (Claude).

2. táblázat: Kutatásból kizárt MI platformok és kizárásuk oka

MI platformok	Kiválasztással kapcsolatos komment	Elérhetőség
you.com Smart	bizonyos számú kérdés után nem tudott válaszolni a kérdésekre	https://you.com/
Chatsonic Ingyenes verzió	havonta csak 25 kérdést lehet feltenni	https://app.writesonic.com/
Perplexity.ai ingyenes verzió	5 kérdés engedélyezett 4 óránként	https://www.perplexity.ai/
HuggingChat	nem volt elérhető a kutatás ideje alatt	https://huggingface.co/
Jasper	csak fizetős elérhetőség	https://www.jasper.ai/
Claude	Nem elérhető Magyarországon	https://claude.ai/
Socratic	csak applikáció formájában létezik	https://socratic.org/
AskYourpdf	csak angolul tud válaszolni, nem érti a komplexebb kérdéseket	https://askyourpdf.com/
pdf.ai	kérdés szám limitáció 1000 karakter	https://pdf.ai/

Forrás: Saját kutatás alapján saját szerkesztés (2024)

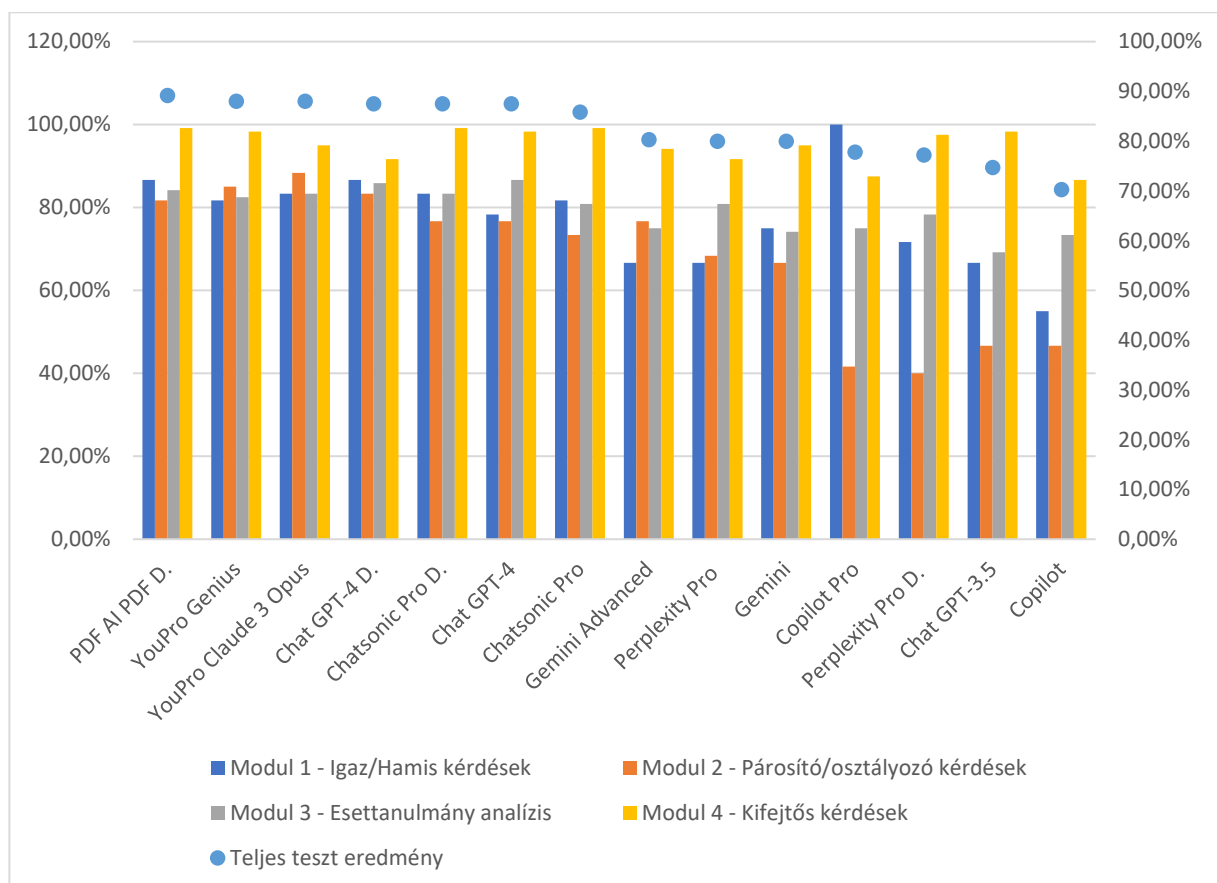
A kutatás során több módszertani kihívással szembesültünk. A komplex kérdések gyakran hibás vagy hiányos válaszokat eredményeztek, ami indokolta a promptok újrafogalmazását, és rávilágított arra, hogy az emberi értékelő szerepe kulcsfontosságú a vizsgálat következetességének

biztosításában. Azonos promptokra eltérő válaszok születtek ugyanattól az MI-től, ami a generatív rendszerek inkonzisztenciáját mutatta. Egyes platformok (például Copilot, Gemini) nem tudtak PDF-dokumentumokat feldolgozni, csak képfájlokat, ami korlátozta a vizsgálat szimulációs értékét, míg mások, mint az OpenAI ChatGPT-4, kifejezetten jól teljesítettek ezen a téren, előnyt biztosítva a tudásbázis integráció és információfeldolgozás szempontjából. Emellett előfordultak szerverterhelésből fakadó elérhetőségi problémák is, amelyek a gyakorlatban a vállalati onboarding folyamat megszakadásának feleltethetők meg, rávilágítva arra, hogy az MI-alapú tudástranszfer nemcsak tartalmi, hanem technológiai megbízhatósági kihívásokkal is szembesül.

Eredmények

A kutatás négy típusú kérdéscsoport - igaz-hamis (Modul 1), párosítás (Modul 2), esettanulmány (Modul 3) és kifejtős kérdések (Modul 4) - teljesítményét vizsgálta különböző mesterséges intelligencia (MI) platformokon. A vizsgálat három kategóriába sorolta a platformokat: ingyenes eszközök, fizetős eszközök és dokumentum feltöltése utáni eredmények.

A válaszok pontossága során azt vizsgáljuk, hogy a megoldókulcs alapján mennyire helyes válaszokat adott vissza az MI. Az MI platformok összesen a feltett 3360 kérdés 82,4%-át tudták jól megválaszolni.



1. ábra: MI platform kutatás eredmények kérdés modulonként (százalékban)

Forrás: Saját kutatás alapján saját szerkesztés (2024)

Összességében azt láthatjuk, hogy a PDF AI PDF D. platformnak sikerült a legjobb eredményt elérnie (89%) a többi platform között, viszont nagyon kis eltéréssel. Az első 7 platform 85% feletti helyes válaszokat adott (PDF AI PDF D., YouPro Genius, YouPro Claude 3 Opus, ChatGPT-4 D. Chatsonic Pro D., ChatGPT-4, Chatsonic Pro). Az eredményekből is látszik, hogy a fizetős platformok jobban teljesítenek az ingyenes verzióknál, de a dokumentum felöltése nem javította szignifikánsan az eredményeket.

Statisztikai elemzés

Az egyirányú varianciaanalízis (ANOVA) kimutatta, hogy szignifikáns különbség van az egyes AI platformok teljesítménye között ($F(13, 3346) = 10.91, p < 0.001$). Ez azt jelzi, hogy a platformok eltérő mértékben voltak képesek támogatni a vizsgafeladatok megoldását, vagyis nem homogén a teljesítményük a tudástranszfer hatékonyságát tekintve.

A különbségek pontos azonosítása érdekében Tukey-féle post-hoc tesztet végeztünk. Az eredmények azt mutatják, hogy több platform szignifikánsan jobban teljesített, mint a ChatGPT-3.5, amely az egyik leggyengébb eredményt produkálta. Különösen a ChatGPT-4 dokumentumfeltöltős verziója ($p < 0.001$), a PDF AI PDF D. ($p < 0.001$), valamint a YouPro Claude 3 Opus és YouPro Genius ($p < 0.001$) szignifikánsan magasabb százalékos teljesítményt értek el.

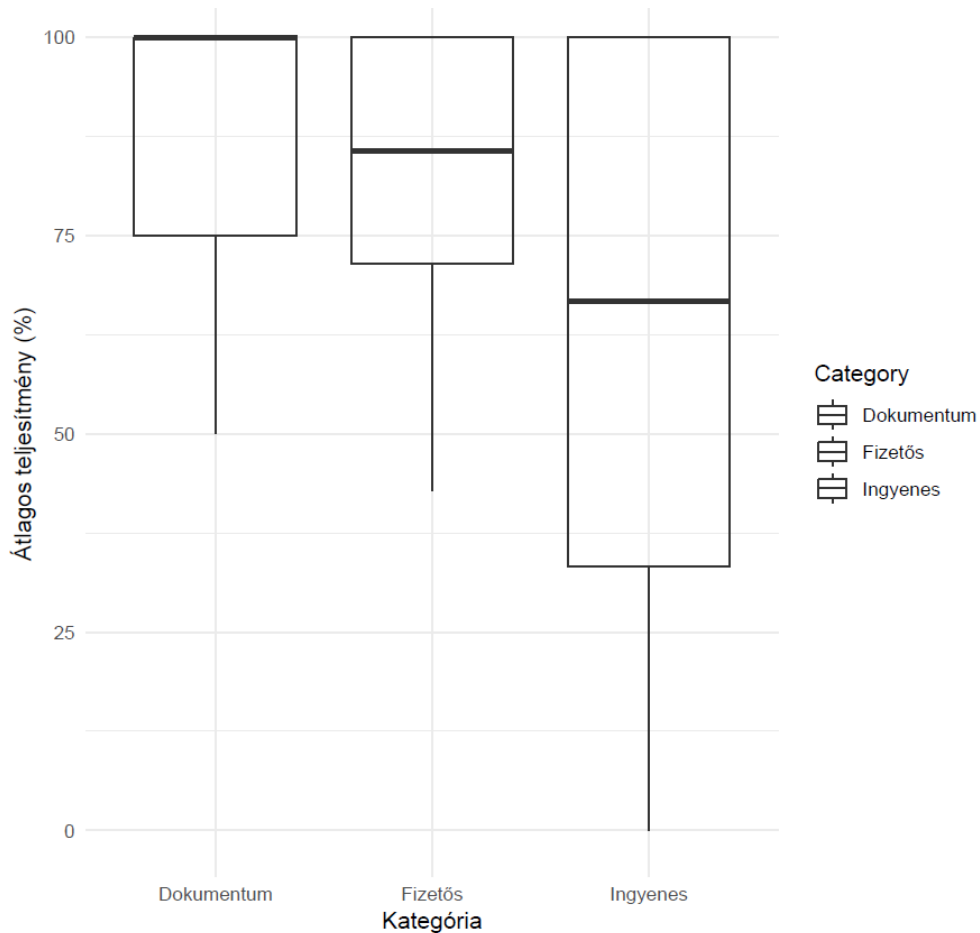
Ezzel szemben a Copilot és Copilot Pro platformok szignifikánsan alacsonyabb teljesítményt mutattak, mint például a ChatGPT-4 vagy a dokumentumfeltöltős verziók ($p < 0.001$). A Perplexity Pro D. platform teljesítménye szintén szignifikánsan elmaradt a PDF AI PDF D.-hez képest ($p = 0.0008$).

Az ábra a Tukey post-hoc teszt konfidenciaintervallumait szemlélteti (95% family-wise CI), amelyen jól látható, hogy több összehasonlításnál a 0-tól eltérő konfidenciaintervallum igazolja a szignifikáns különbségeket.

MI kategóriák összehasonlítása

Az ingyenes, fizetős és dokumentumfeltöltős AI-platformok teljesítményét összehasonlító varianciaanalízis (ANOVA) szignifikáns különbségeket mutatott a kategóriák között ($F(2, 3357) = 21.26, p < 0.001$). A post hoc Tukey-próba eredményei szerint az ingyenes platformok teljesítménye szignifikánsan alacsonyabb volt mind a fizetős (átlagkülönbség = -9.80% , $p < 0.001$), mind a dokumentumfeltöltős platformokhoz képest (átlagkülönbség = -12.52% , $p < 0.001$). Ugyanakkor a fizetős és a dokumentumalapú rendszerek között nem volt kimutatható statisztikailag szignifikáns eltérés (átlagkülönbség = -2.72% , $p = 0.225$).

Az átlagteljesítményt bemutató boxplot (1. ábra) alapján az ingyenes platformok mediánja alacsonyabb, eloszlásuk szélesebb, ami nagyobb teljesítményingadozásra utal. A fizetős és a dokumentumfeltöltős kategóriák magasabb és stabilabb teljesítményt mutattak, mediánjuk közel esik egymáshoz, és a felső kvartilisük is szűkebb tartományban helyezkedik el.



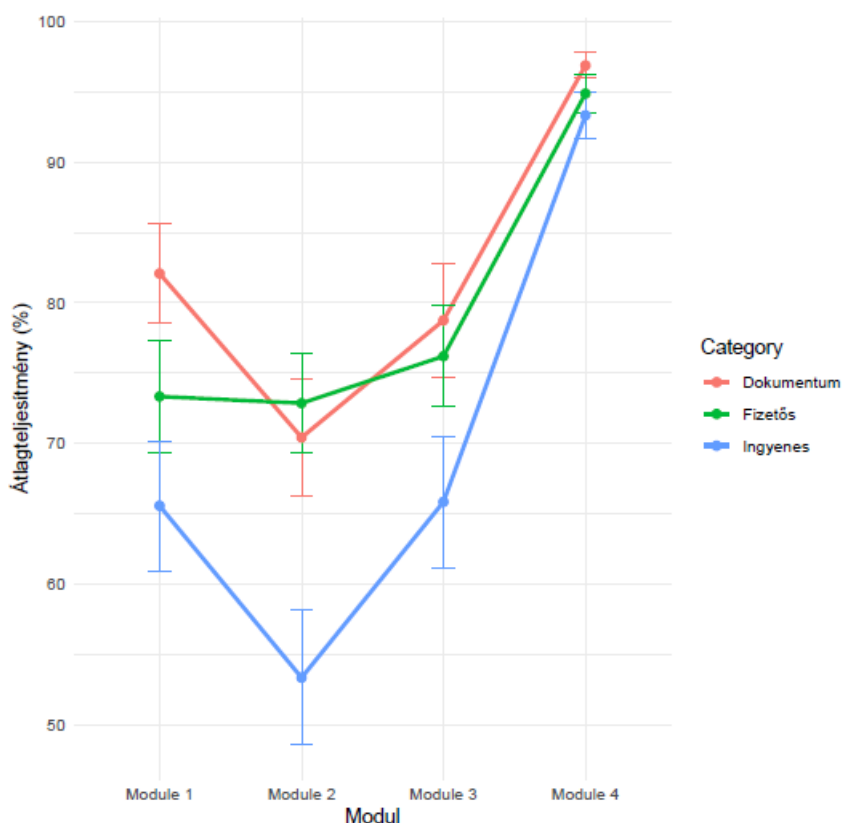
2. ábra: MI platform kutatás: átlagos teljesítmény kategóriák szerint (százalékban)

Forrás: Saját kutatás alapján saját szerkesztés (2024)

Interakciós hatások vizsgálata

A két tényezős varianciaanalízis (Category \times Module) eredményei alapján mind a kategória ($F(2,708) = 13.09$, $p < 0.001$), mind a modul típusa ($F(3,708) = 36.18$, $p < 0.001$) szignifikáns főhatást mutatott. Ez azt jelzi, hogy az MI-platformok teljesítménye egyrészt kategóriánként (ingyenes, fizetős, dokumentumalapú), másrészt a kérdésmodulok típusa szerint is érdemben eltért. Ezzel szemben a két tényező közötti interakció nem bizonyult szignifikánsnak ($F(6,708) = 1.46$, $p = 0.189$), vagyis a modulok szerinti eltérések iránya és mértéke alapvetően hasonló volt a különböző platformkategóriákban.

Az ábra alapján ugyanakkor jól látható, hogy bizonyos modulok esetében nagyobb különbség mutatkozott a platformkategóriák között. Például az igaz-hamis kérdéseket tartalmazó Modul 1-ben a dokumentumalapú rendszerek (átlag $\sim 82\%$) jobban teljesítettek, mint az ingyenes eszközök ($\sim 66\%$), a fizetős rendszerek pedig köztes eredményt mutattak ($\sim 73\%$). A párosítási feladatoknál (Modul 2) a különbségek tovább erősödtek: az ingyenes platformok teljesítménye látványosan visszaesett (medián $\sim 54\%$), míg a fizetős és dokumentumalapú rendszerek stabilabb eredményt adtak ($\sim 70\text{--}72\%$). Az esettanulmányoknál (Modul 3) ismét kiegyenlítettebb teljesítmény figyelhető meg, ugyanakkor az ingyenes kategória továbbra is alacsonyabb értékeket mutatott. A kifejtős kérdéseknél (Modul 4) minden kategória teljesítménye jelentősen megemelkedett, és közel $90\text{--}96\%$ -os átlagot ért el, a különbségek itt már elhanyagolhatók.



3. ábra: Interakciós hatás vizsgálat (százalékban)

Forrás: Saját kutatás alapján saját szerkesztés (2024)

Értelmezés

Az eredmények azt mutatják, hogy a dokumentumalapú és a fizetős platformok konzisztensen magasabb és stabilabb teljesítményt nyújtanak, különösen az olyan feladatoknál, amelyek összetettebb kapcsolatteremtést igényelnek (pl. Modul 2: párosítás, Modul 3: esettanulmány). Az ingyenes platformok ezzel szemben jóval nagyobb ingadozást mutattak, ami azt jelzi, hogy kevésbé megbízhatóak a komplexebb tudástransfer-feladatokban. A Modul 4-ben (kifejtős kérdések) tapasztalt, kategóriák közötti különbségek eltűnése arra utal, hogy amikor a feladat nagy mennyiségű szöveges kimenetet igényel, minden típusú rendszer képes közel azonos szinten teljesíteni. Ez vélhetően annak tudható be, hogy a legtöbb MI-platform alapvetően szövegkimenetek generálására lett optimalizálva.

Összességében a statisztikai eredmények megerősítik, hogy bár az interakció nem szignifikáns, az egyes modulok jellegzetességei erősen befolyásolják a különböző kategóriájú rendszerek teljesítményét. Ez alapján a kutatás rámutat arra, hogy a tudástransfer hatékonyságának vizsgálatakor nem elegendő pusztán a platformkategóriákat összevetni, hanem a feladat típusa és kognitív igényessége is meghatározó tényező.

Hipotézisek elfogadása/megcáfolása

1. Hipotézis: Feltételezzük, hogy az ingyenes MI platformok kevesebb helyes választ adnak, mint a fizetős változataik, és ezek válaszási képessége korlátozottabb, például egyszerre csak korlátozott számú kérdésre képesek válaszolni, míg a fizetős változatoknál ez a korlátozás nem áll fenn.

Az eredmények összhangban vannak a feltevésével, miszerint az ingyenes platformok gyengébben teljesítenek a fizetősöknél. A kategória-szintű egyirányú ANOVA szignifikáns főhatást mutatott ($F(2, 3357)=21.26, p<0.001$), a Tukey-összehasonlítás pedig egyértelműen jelezte, hogy az ingyenes kategória átlagteljesítménye szignifikánsan alacsonyabb mind a fizetős (átl. különbség -9.80 százalékpont, $p<0.001$), mind a dokumentumalapú rendszerekhez képest (-12.52 százalékpont, $p<0.001$). A fizetős és dokumentumos csoport között ugyanakkor nem találtunk szignifikáns különbséget (-2.72 százalékpont, $p=0.225$). Ez a mintázat modulonként is konzisztensen megjelent: a különbségek a strukturáltabb feladattípusoknál (igaz-hamis, párosítás) voltak a legnagyobbak, míg a szabad szövegű (kifejtős) feladatoknál elenyészőre csökkentek.

A hipotézis második része – miszerint az ingyenes eszközök válaszadási kapacitás-korlátai (kérdésszám, tempó) önmagukban rontják a teljesítményt – közvetlenül nem volt mérhető a jelen adathalmazból, mert nem gyűjtöttünk platformszintű throttling/limit-időbélyeg adatokat. A mintázat azonban összhangban áll ezzel a mechanizmussal: az ingyenes eszközök nagyobb szórása és alacsonyabb mediánja arra utal, hogy gyakrabban ütköznek erőforrás- és funkciókorlátokba (pl. rövidebb kontextusablak, gyengébb modellverzió, kérdéslimit), ami főként a pontos tényfelidézést és a fogalmi kapcsolatok következetes kezelését igénylő feladatoknál (Modul 1–2) jelent hátrányt. Összességében H1-et elfogadjuk: a fizetős platformok statisztikailag szignifikánsan és gyakorlatilag is jobban teljesítenek az ingyeneseknél a vizsgált tudástranszfer-feladatok többségében; a kapacitáskorlátokra vonatkozó magyarázat plauzibilis, de további (napló- és terhelési) adatokkal lenne érdemes külön igazolni.

2. Hipotézis: Azok az MI rendszerek, amelyek dokumentumok betöltésére képesek, szignifikánsan jobb válaszadási teljesítményt mutatnak majd, különösen az esettanulmányokkal kapcsolatos kérdések esetében, ahol nem csak elméleti tudásra, hanem annak gyakorlati alkalmazására is szükség van.

Az eredmények részben támasztják alá ezt a hipotézist. Az ANOVA és a Tukey-teszt eredményei szerint a dokumentumalapú rendszerek teljesítménye szignifikánsan jobb az ingyenes platformokénál (-12.52 százalékpont, $p<0.001$), ugyanakkor nem különbözik szignifikánsan a fizetős eszközökétől (-2.72 százalékpont, $p=0.225$). A kéttényezős ANOVA (Kategória \times Modul) alapján az interakció nem bizonyult szignifikánsnak ($F(6,708)=1.46, p=0.189$), vagyis statisztikai értelemben nem igazolható, hogy a dokumentumfeltöltés kifejezetten az esettanulmányoknál nyújtana többletet. Ugyanakkor a trendek következetesen kedvezőek: Modul 1-ben és Modul 3-ban a dokumentumalapú eszközök átlagai magasabbak ($\sim 82\%$ és $\sim 80\%$), és több páronkénti összevetésben is előnyt mutattak (pl. a PDF AI PDF D. szignifikánsan jobb a Perplexity Pro D.-nél, $p=0.0008$).

A hipotézis azon része, miszerint a dokumentum-betöltés különösen az esettanulmányoknál eredményezne szignifikánsan jobb teljesítményt, nem igazolódott. Az eredmények inkább arra utalnak, hogy a dokumentumalapú rendszerek általános előnyt képviselnek az ingyenesekkel szemben, de a fizetős platformokhoz képest nem mutatnak egyértelmű fölényt.

3. Hipotézis: Azt várjuk, hogy az általános tudásra épülő kifejtős kérdésekre majdnem minden MI képes lesz helyes választ adni, mivel ezek az információk könnyen hozzáférhetőek az adatbázisukban, de az igaz-hamis állítások és osztályozós feladatoknál jelentős eltérések lesznek a különböző platformok teljesítménye között. Ezen kívül előfordulhatnak formai helyességgel kapcsolatos problémák is, például, ha nem megfelelően van promptolva a feladat, ami eltérő eredményeket eredményezhet azonos prompt esetén.

Az eredmények teljes mértékben alátámasztják a hipotézist, miszerint a kifejtős kérdésekben (Modul 4) csak csekély különbségek tapasztalhatók, míg az igaz-hamis és párosítási feladatoknál a platformok közötti eltérések jelentősek. A kéttényezős ANOVA erős modul-főhatást mutatott

($F(3,708)=36.18$, $p<0.001$): a Modul 4 átlagos teljesítménye minden kategóriában kiemelkedően magas volt (~90–96%), és az eltérések statisztikailag elhanyagolhatónak bizonyultak. Ezzel szemben a Modul 1 és Modul 2 esetében a különbségek szignifikánsak: az ingyenes platformok rendre gyengébben teljesítettek (pl. Modul 2-ben a medián mindössze ~54%), míg a fizetős és dokumentumalapú rendszerek stabilan a ~70–72%-os tartományban mozogtak. Az egyirányú ANOVA platformszinten is szignifikáns eltéréseket mutatott ($F(13,3346)=10.91$, $p<0.001$), amelyeket a Tukey-teszt is megerősített, például a Copilot és Copilot Pro szignifikánsan gyengébb eredményeit a ChatGPT-4-hez és több dokumentumalapú konfigurációhoz képest ($p<0.001$).

Ezek az összefüggések jól illeszkednek a generatív modellek működésének természetéhez. A kifejtős kérdések esetében, ahol a szövegenerálás és az érvelő kimenet előállítás a fő feladat, a platformok teljesítménye konvergál, hiszen mindegyik modell alapvetően erre az outputtípusra optimalizált. A tényellenőrzést vagy strukturált logikai kapcsolatok felismerését igénylő kérdéstípusokban azonban a rendszer- és konfigurációs különbségek sokkal inkább megmutatkoznak. Ez azt jelzi, hogy míg a kifejtős feladatok alkalmasak lehetnek az MI szövegalkotási kapacitásának összevetésére, addig a tudástranszfer megbízhatóságának értékeléséhez érdemes az igaz–hamis és párosító modulokra támaszkodni, ahol a teljesítménykülönbségek jobban differenciálnak az eszközök között.

Következtetések és javaslatok

Összegzés. A vizsgálatunk azt mutatja, hogy a generatív MI alkalmas eszköz lehet az explicit tudás szervezeti közvetítésére: a platformok összesen a 3360 kérdés 82,4%-ára adtak helyes választ. A teljesítmény ugyanakkor nem homogén. Kategóriaszinten az ingyenes rendszerek szignifikánsan gyengébben teljesítettek a fizetős és a dokumentumalapú konfigurációkhoz képest (ANOVA/Tukey), miközben a dokumentum-betöltés a legtöbb modulban általános előnyt jelentett az ingyenesekhez viszonyítva, de nem eredményezett szignifikáns többletet a fizetős rendszerekhez képest. Modulonként erős főhatást találtunk: a kifejtős feladatokban (Modul 4) a rendszerek teljesítménye konvergál (~90–96%), míg a tényalapú/strukturált feladatokban (igaz–hamis, párosítás) látványosak a különbségek. Eredményeink illeszkednek ahhoz az irodalmi képhez, amely szerint a generatív modellek kiválóak szöveg- és összefoglaló generálásban (Pedersen, 2023; Toma et al., 2023), de a precíz ténykezelés és a strukturált relációk következetes kezelése érzékeny a rendszer- és konfigurációs különbségekre, valamint a promptolás minőségére (Sivakumar et al., 2024).

A tudástranszfer keretrendszerében a generatív MI különösen az explicit tudás externalizációját és (re)internalizációját támogatja: dokumentált ismeretekből feladatra szabott, konzisztens narratívákat képez (Chikán, 2008). Ugyanakkor a hitelesség kulcsfontosságú: a generált tartalom azonosíthatósága és ellenőrizhetősége nélkül nő a tévedések és a dezinformáció kockázata (Najee-Ullah et al., 2022). A chatbot-interakció mint közeg képes személyre szabott támogatást adni (Adamopoulou & Moussiades, 2020; Chan & Hu, 2023), de a rendszerszintű megbízhatóság és governance nélkülözhetetlen (Aronsson et al., 2021).

A gyakorlati következtetések szerint a kritikus onboarding- és tudástár-folyamatoknál érdemes fizetős vagy dokumentumalapú konfigurációt választani, mert ezek stabilabban teljesítenek, különösen a kapcsolat- és kontextusérzékeny feladatokban (párosítás, esettanulmány). Ennek feltétele egy „gold standard” vállalati tudásbázis kialakítása és karbantartása – jóváhagyott szabályzatokkal, folyamatleírásokkal, GYIK-kal és esettárral –, amelyet rendszeresen be kell tölteni

az MI-rendszerekbe, és verziózott hivatkozásokkal kell visszakerülni a kimenetekben a forrásellenőrizhetőség érdekében (Najee-Ullah et al., 2022). A promptolást érdemes formalizálni (feladatléírás, szerepkör, kontextus, elvárt forma) és mintapromptokkal támogatni a felhasználókat (Sivakumar et al., 2024). A minőségbiztosítás része legyen a kötelező hivatkozáskérés, a források visszaellenőrzése, a vitatható válaszok megjelölése és a „hallucinációs” minták gyűjtése a későbbi tréninghez. A tacit és explicit tudás átadását célszerű kombinálni: a generatív MI mikrotananyagokkal és helyzetfüggő sűgőval kísérje a mentorálást/peer-reviewt, míg a humán szereplők adják a normatív és kulturális kontextust (Chikán, 2008; Chan & Hu, 2023). Üzemeltetési oldalról indokolt a rendelkezésre állási/terhelési incidensek számlálása és egy előre kialakított „fallback” (alternatív platform, offline tudástár), mert az elérhetőségi zavarok a gyakorlatban onboarding-megszakításként jelennek meg.

A bevezetést célszerű kisléptékű pilottal indítani egy kijelölt üzleti területen (például ügyfélszolgálat vagy marketing), dokumentumalapú vagy fizető MI-konfigurációval. Már a pilotban rögzítsünk mérőszámokat: kérdésszintű pontosság, konzisztencia ismételt kérdésnél, hivatkozott források aránya, válaszüő és kritikus hibák száma. Ezzel párhuzamosan épüljön fel a prompt- és sablontár, valamint a felhasználói tréning. A működtetést folyamatos értékelés és visszacsatolás kísérje: havi „vizsga” a friss tudástár-tartalmakból, hibakatalógus és javítóintézkedések karbantartása.

A kutatás korlátai között szerepel, hogy méréseink szimulált vizsgakörnyezetben történtek (nem valós onboarding), a platformok 2024. áprilisi állapotát vizsgáltuk és magyar nyelvi környezetben értékeltünk; nem álltak rendelkezésre részletes throttling/limit vagy időbélyeg adatok; a pontozás néhány elemében szubjektív komponensek is megjelentek. Mindez konzervatív értelmezést indokol, különös tekintettel a generált tartalmak pontosságára és a dezinformációs kockázatokra (Faust et al., 2023).

További kutatási irányként javasoltak a valós onboarding-kísérletek idő- és minőségi kimenőkkel (idő-a-kompetenciáig, hibaarány, tudásmegtartás), megbízhatósági mutatók (például Cronbach-alfa modulonként) és vegyes hatású modellek alkalmazása a platform- és kérdésszintű variancia szétválasztására, valamint az interakciók finomabb vizsgálata (feladattípus \times tudásbetöltés). Érdemes mérni a hallucináció-detektálás és a forrásellenőrzési protokollok hatását, továbbá vizsgálni a felhasználói élményt és elfogadottságot, különös tekintettel a személyre szabott támogatás motivációs hatásaira (Chan & Hu, 2023).

Összegzésként: a generatív MI – megfelelő dokumentum-integrációval, prompt-fegyelemmel és minőségbiztosítással – érdemben növelheti az explicit tudás elérésének és közvetítésének hatékonyságát. A tacit dimenziók miatt nem helyettesíti, hanem kiegészíti a humán tudásátadást; az optimális vállalati megoldás a hibrid (ember+MI) tudásmenedzsment, amely a szakirodalom által leírt előnyöket (Pedersen, 2023; Toma et al., 2023) a kockázatkezeléssel és a governance-szabályokkal hozza egyensúlyba (Najee-Ullah et al., 2022; Aronsson et al., 2021).

Hivatkozott források

Luz, A. – Olaoye, G. O. (2024): Artificial Intelligence and Employee Experience: Leveraging Technology for Personalization. *EasyChair*, 13219. <https://easychair.org/publications/preprint/27cs/open>

Pedersen, I. (2023): The rise of generative AI and enculturating AI writing in postsecondary education. *Front. Artif. Intell.* 6:1259407. <http://dx.doi.org/10.3389/frai.2023.1259407>

- Adamopoulou, E. – Moussiades, L. (2020): Chatbots: History, Technology, and Applications. *Machine Learning with Applications*, 2, 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Sivakumar, A. – Gelman, B. – Simmons, R. (2024): Standardized nomenclature for litigational legal prompting in generative language models. *Discover Artificial Intelligence*, 4(21). <https://doi.org/10.1007/s44163-024-00108-5>
- Toma, I. – Umbrich, J. – Theil, M. – Sas, A. – Reinwald, W. H. – Theil, A. – Wahler, A. – ŞİMŞEK, U. – Fensel, D. (2023): Towards the digital transformation of Distribution System Operators using Knowledge Graphs and Conversational AI. *27th International Conference on Electricity Distribution (CIRED 2023)*, Rome, Italy, 2023, pp. 306–310. <https://doi.org/10.1049/icp.2023.0298>
- Chikán Attila (2008): *Vállalatgazdaságtan*, 4. átdolgozott, bővített kiadás, Budapest, AULA Kiadó, 332–334 p.
- Najee-Ullah, A. – Landeros, L. – Balytskyi, Y. – Chang, S.-Y. (2022): Towards Detection of AI-Generated Texts and Misinformation. In Parkin, S. – Viganò, L. (Eds.), *Socio-Technical Aspects in Security, Lecture Notes in Computer Science*, vol 13176. Springer, Cham. https://doi.org/10.1007/978-3-031-10183-0_13
- Chan, C.K.Y. – Hu, W. (2023): Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education. ResearchGate. <https://doi.org/10.1186/s41239-023-00411-8>
- Aronsson, J. – Lu, P. – Strüber, D. – Berger, T. (2021): A Maturity Assessment Framework for Con-versational AI Development Platforms. In *The 36th ACM/SIGAPP Symposium on Applied Computing (SAC '21), March 22–26, 2021*. Virtual Event, Republic of Korea. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3412841.3442046>
- Faust, A. – Dröge, M. – Odebrecht, C. (2023): Assessment of AI literacy Development and testing of a cus-tomizable set of items. *Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft für Informatik (GI)*, P-337, 341–346

Szerzők

Tétényi Bálint

ORCID: 0009-0004-7000-4979

MSc

Global marketing excellence manager

Richter Gedeon Nyrt.

tetenyi.balint@gmail.com

Dr. Kovács Stefan

ORCID: 0000-0002-2644-8781

PhD

Egyetemi adjunktus

Budapesti Műszaki és Gazdaságtudományi Egyetem, Gazdaság- és Társadalomtudományi Kar,
Menedzsment és Vállalkozásgazdaságtan Tanszék

kovacs.stefan@gtk.bme.hu

Csapó Anna
MSc
Senior Business Analyst & Digital Project Specialist
Richter Gedeon Nyrt.
csaannaa@gmail.com

Kádár-Buksa Márk
Szabadúszó
kadarmark@gmail.com

A műre a Creative Commons 4.0 standard licenc alábbi típusa vonatkozik: [CC-BY-NC-ND-4.0.](https://creativecommons.org/licenses/by-nc-nd/4.0/)

