

A MESTERSÉGES INTELLIGENCIA HATÁSA A MÉRÉS-ÉRTÉKELÉSRE

MOLNÁR GYÖNGYVÉR^{a,b}

^aSZTE Neveléstudományi Intézet, Szeged

^bMTA-SZTE Digitális Tanulási Technológiák Kutatócsoport, Szeged

Beérkezett: 2023. augusztus 13., elfogadva: 2024. január 8.

Minden új technológia megjelenése lehetőséget ad arra, hogy átgondoljuk az oktatás céljait, hatékonyságát. A mesterséges intelligencia mérés-értékelésben való alkalmazása nemcsak új lehetőségeket kínál, hanem komoly kihívások elé is állítja az oktatás szereplőit. A mesterségesintelligencia-alapú értékelési technikákban, melyek lehetővé teszik nagy adathalmazok gyors elemzését és a személyre szabott visszacsatolás megvalósítását, benne van az a lehetőség, hogy a hagyományos értékelési módszerek átalakításával jelentős mértékben fokozzák a tanulási folyamat hatékonyságát és emeljék a tanítás minőségét. Autentikus, megbízható és valid adatok nélkül azonban az adatokon, bizonyítékokon alapuló fejlesztés kudarcra van ítélve. A mesterséges intelligencia mérés-értékelésben történő alkalmazása nem csupán technológiai vagy infrastrukturális, sokkal inkább pedagógiai és módszertani kihívás.

Kulcsszavak: mesterséges intelligencia, mérés-értékelés, diagnosztikus értékelés, viselkedésadatok elemzése, személyre szabott oktatás

THE IMPACT OF ARTIFICIAL INTELLIGENCE ON ASSESSMENT

Every time a new technology emerges, it provides an opportunity to reconsider the purpose of education. Advanced AI technology offers new opportunities and new challenges in education, specifically in educational assessment. AI-based assessment has the potential to transform traditional assessment by processing a large amount of data and providing learners with personalized feedback, thereby enhancing the efficacy of the learning process and improving teaching quality through a re-examination of assessment activities. However, without authentic, reliable and valid data, any assessment- and data-based improvement would be doomed to failure. Using artificial intelligence in assessment is not merely a technological or infrastructural matter; it is first and foremost a pedagogical and methodological challenge.

Keywords: artificial intelligence, assessment and evaluation, diagnostic assessment, analyses of behavioural data, personalised education

Levelező szerző: Molnár Gyöngyvér DSc, SZTE Neveléstudományi Intézet, 6723 Szeged, Petőfi S. sgt. 32–34.
E-mail: gymolnar@edpsy.u-szeged.hu

A technológia oktatási alkalmazása nem újkeletű. Leggyakoribb felhasználási módjai az információ gyorsabb hozzáférését és áramlását, motiválóbb tanulási környezetek kialakítását, illetve a távoktatás megvalósítását támogatják. Minden egyes alkalommal, amikor új technológia jelenik meg, eltérő, néha szélsőséges vélemények is megfogalmazódnak az adott technológia oktatási hasznossága vagy épp hátránya vonatkozásában. Például, amikor széles tömegek számára elérhetővé vált az internet és elindultak az online képzések, számos oktató tartott attól, hogy ezen lehetőség kiszorítja a hagyományos oktatási formát (Liu–Baucham 2023). Helyette mára az online oktatás a hagyományos oktatási környezet fontos és szerves részévé vált, megdőntve azt a 20. századi nézetet, miszerint értelmes tanulás csak osztálytermi környezetben történhet.

Világszerte hatalmas vitákat és vegyes reakciókat váltott ki a 2022 novemberében széles tömegek számára elérhetővé vált ChatGPT oktatási alkalmazása. Néhány iskola, illetve oktatási rendszer azonnali betiltotta használatát (Jimenez 2023), mások a ChatGPT által generált szöveges tartalom felismerését támogató ellenőrző alkalmazásokat fejlesztettek és alkalmaztak (DetectGPT – Miller 2023 vagy GPTZero – Svrluga 2023), míg szintén mások ösztönözték a ChatGPT tanórai munkában kreatív módon megjelenő használatát (The Associated Press 2023).

Az elmúlt két évtizedben igen jelentős mennyiségű kutatási eredmény, publikáció született a technológiával támogatott oktatás témakörén belül (Adesope–Rud 2019), jelentősen növelve a kutatásokban már elérhető vs. a gyakorlatban alkalmazott eszközök, módszerek közötti különbségek mértékét. Ez nincs máshogy a mérés-értékelés területén sem. A kutatási eredmények alapján megfogalmazható, hogy a technológia iskolai, oktatási alkalmazásának tiltása sohasem működött, ugyanakkor használata annál hatékonyabbnak bizonyult, minél közelebb állt az eszköz segítségével kidolgozott oktatási anyag a valósághoz és minél jelentősebb volt a fejlesztő/értékelési eszköz pedagógiai beágyazottsága (Molnár–Turcsányi–Szabó–Kárpáti 2020).

A forradalmian új technológiák megjelenése minden egyes alkalommal lehetőséget ad arra, hogy átgondoljuk az oktatás célját és tartalmát, kritikusan felülvizsgáljuk az alkalmazott eszközöket, módszereket, lehetőségeket, az oktatási rendszer minőségének és hatékonyságának egészét. Az oktatás azon szektorok közé tartozik, amelyeket nemcsak a digitalizáció nagyerejű terjedése, hanem a társadalmi környezet, a munkaerőpiaci elvárások egyéb, gyorsan változó kihívásai is befolyásolnak. A tanulás és az oktatás hatékonyságának növelését célzó kutatások eredményeinek hatására nemzetközi szinten előtérbe került a személyre szabott tanulás és tanítás (smart education) fontossága (Singh–Miah 2020; Molnár–Csapó 2023), illetve ezzel szoros összefüggésben jelentősen megváltozott a mérés-értékelés célja, funkciója, tárgya és szerepe. A nagy téttel bíró szummatív és a kis tétellel bíró diagnosztikus értékelés közötti váltást felerősítették és felgyorsították az úgynevezett vészhelyzeti távolléti oktatással kapcsolatos tapasztalatok, a mérés-értékelés tárgyának, típusának, jellegének újragondolását pedig a mesterséges intelligencia széles körű hozzáférhetővé válása segítette.

A mesterséges intelligencia (MI) mérés-értékelésbeli alkalmazása során lényeges szem előtt tartani, hogy a mesterséges intelligencia – hasonlóan a többi technológiához – „csak” egy új lehetőséget biztosító eszköz, amely megfelelő alkalmazás mellett hozzájárulhat a meglévő oktatási problémák megoldásához, a személyre szabottabb tanulási és értékelési környezetek kialakításához, valamint ezzel összefüggésben a tanulás és az oktatás hatékonyságának növeléséhez. Alkalmazása, használata nem lehet oktatási cél (Molnár 2011), azaz ne ezen eszköz, eszközök határozzák meg a változtatások, fejlesztések irányát, hanem maguk a megoldandó problémák (Molnár–Turcsányi–Szabó–Kárpáti 2019, 2020). A mesterséges intelligencia oktatási alkalmazása nem csak technológiai, infrastrukturális kérdés, sokkal inkább pedagógiai, módszertani kihívás.

A tanulmány keretein belül a technológia, külön kiemelten a mesterséges intelligencia adta olyan új lehetőségek alkalmazásával és adaptálási lehetőségeivel foglalkozunk, melyek adatvezérelt világunkban a mérés-értékelés funkciójának, lehetőségeinek átalakításán keresztül hozzájárulnak és

hozzájárulhatnak az oktatás minőségének fejlesztéséhez. Kiemelten foglalkozunk (1) a mérés-értékelés funkciójának, szerepének változásával, (2) a megváltozott társadalmi igényekre reagáló oktatás alapját biztosító diagnosztikus mérés-értékelés szerepének növekedésével, ezen belül a mesterséges intelligencia adta azon lehetőségek kiemelésével, amelyek a diákok között meglévő és a tanévek előrehaladtával egyre növekedő különbségek csökkentését támogathatják, (3) végül áttekintjük a mesterséges intelligencia adta minőségibb és érvényesebb értékelést megvalósító lehetőségek körét – mind a nagy adathalmazokkal dolgozó kutatás-fejlesztésen alapuló értékelési rendszerekben, mind a közvetlen osztálytermi értékelés esetén.

A technológia rapid fejlődésének hatása a mérés-értékelés funkciójának, szerepének, tárgyának és módszereinek változására

„Ahhoz, hogy jól működő oktatási rendszert hozzunk létre, kifinomult visszacsatoló köröket kell a rendszerbe beépítenünk. [...] Nem mindegy azonban, mértékben, értékelnek a visszacsatoló rendszerek. Egyszerűbben fogalmazva, nem lehet cél az, hogy az iskolák még több erőfeszítéssel, esetleg még nagyobb hatékonysággal termeljenek olyan tudást, amelynek jelentős része használhatatlan. Ehelyett arra kell törekednünk, hogy a tanulók jó minőségű, használható tudással hagyják el az iskolát.” (Csapó 2008: 114.) „Az oktatás [...] hatékonyságának és sikerességének is egyik kulcsa, hogy a diákok azt tanulják, aminek elsajátítására felkészültek, amihez rendelkeznek a szükséges alapokkal” (Molnár–Csapó 2019b: 706), azaz a sikeres tanítás fontos feltétele, hogy ismerjük a diákok azon tudás- és képességterületeken való felkészültségét, melyek meghatározó erővel bírnak tanulási sikerességük tekintetében.

Adatgazdag világunkban csak a magas minőségű, megbízható és érvényes mérés során keletkezett adatok képezhetik a fejlesztések forrását, azok értelmezése egyértelműsíti, hogy hol tartanak a diákok a tanulásban, hova kell még eljutniuk, és a célt hogyan, mely területek fejlesztésével érik el leghatékonyabban. Érvényes, megbízható mérés során keletkezett adatok nélkül az értékeléssel kapcsolatos fejlesztések kudarcra vannak ítélve, azaz a mesterséges intelligencia hatékony oktatási alkalmazásának egyik fontos feltétele a minőségi adat. A nem kellően átgondolt és nem megfelelő szakértelemmel, rosszul tervezett értékelés egyrészt frusztrációt eredményez és negatívan befolyásolja a diákok tanulását, illetve tanulási eredményeit (Liu–Baucham 2023), másrészt hamis és nem releváns forrását képezi az oktatási fejlesztéseknek.

Az oktatás területén történt első technológiai alkalmazások a mérés-értékeléshez köthetők. Közel száz évvel ezelőtt történt, hogy Sidney Pressey (1927) kidolgozta, majd felsőoktatási kontextusban alkalmazta a feleletválasztós feladatok azonnali javítását és az eredmény visszacsatolását megvalósító oktatógépet. A technológialapú mérés-értékelés fejlődésének következő, igen jelentős mérföldkővét három évtizeddel ezelőtt a számítógépek, a www és az internet indulása, majd széles körű elterjedése hozta magával, aminek hatására új irányt vettek a vonatkozó kutatások (Greenhow–Robelia–Hughes 2009). Elindult a legsokoldalúbb technológiai eszköz, a számítógép tesztelési folyamatba történő integrálása, és megjelentek a számítógépalapú teszteléssel kapcsolatos kutatások. Az ezredforduló környékén pedig már elkezdődött a nemzetközi nagymintás mérések technológiaalapúvá átalakítása is. Ezzel párhuzamosan egyrészt – a kapcsolatos kutatások számának jelentős növekedése következtében – jelentős mértékben változtak a számítógépalapú tesztelés lehetőségei, céljai, folyamatosan előtérbe kerültek a korábban kihasználatlanul maradt lehetőségek, másrészt megváltozott a társadalomban értékesnek számító tudás típusa, új, korábban nem vizsgált, innovatív feladattípusok alkalmazását igénylő tudás- és képességterületek mérését kellett megoldani (pl. interaktív vagy kollaboratív, együttműködésen alapuló problémamegoldó képesség, IKT-műveltség).

Míg a számítógépalapú tesztelessel kapcsolatos fejlesztések első húsz évében az egyre autentikusabb, életszerűbb, objektívabb, validabb és megbízhatóbb, azaz egyre jobb minőségű szummatív tesztek fejlesztésén volt a hangsúly, ma már nem kérdés, hogy magas minőségű számítógépalapú tesztek tudunk előállítani (l. pl. PISA-tesztek). Míg a kutatások korai stádiumában nagy hangsúly helyeződött a hagyományos (szemtől szembeni, illetve papíralapú) és technológiaalapú teszteleési eredmények összevetésére, ma már nem kérdés, hogy még az óvodás gyerekek számára sem jelent problémát az érintőképernyős eszközökön történő válaszadás, sokkal fontosabb kérdéssé vált, hogy a technológia adta lehetőségeket hogyan lehet a tanulás hatékonyságának növelése érdekében maximalizálni. Míg korábban a „one fits for all” megközelítés fényében az összehasonlíthatóság érdekében minden egyes diák ugyanazt a tesztet oldotta meg függetlenül tudás- és képességszintjétől, addig ma már – jellemzően a mesterséges intelligencia adta nagy adathalmazok elemzésén alapuló algoritmusok használatával – a képességszintek összehasonlíthatósága mellett lehetőség van a tudás- és képességszinthez igazodó, úgynevezett adaptív tesztek alkalmazására (Csapó–Molnár–R. Tóth 2008; Magyar 2012). Míg korábban a tesztelt személyek (jellemzően diákok) által adott válaszok képezték kizárólagosan az értékelés alapját, addig ma már a kontextuális, azaz a diákok tesztelés során mutatott viselkedési adatainak – jellemzően mesterséges intelligenciához sorolható – mintázatkereső algoritmusokkal történő elemzése (pl. feladatmegoldással töltött idő, a feladatokkal történt interakció mennyisége és milyensége) jelentősen árnyalja és pontosítja a hagyományos teszteredményeket, amelyekről a diákok személyre szóló visszacsatolást kaphatnak.

Mindezen lehetőségek, azaz egyrészt a technológiai eszközök, adatrögzítési módszerek és elemzési repertoár fejlettségi szintje (beleértve a mesterséges intelligencia adta lehetőségeket is), másrészt a tanulás és tanítás minőségének javítását célzó kutatások eredményeinek hatására az elmúlt tíz évben elkezdett megváltozni a mérés-értékelés célja és funkciója (Redecker–Johannessen 2013). Bár továbbra is elvárás maradt, sőt a nagy adathalmazok világában egyre fontosabbá vált a megbízható és érvényes adatokat biztosító magas minőségű tesztelés megvalósítása, de az azonnali visszacsatolás lehetősége a tanulási folyamatot záró nagy tétellel bíró szummatív értékelés helyett/mellett fokozatosan előtérbe helyezte a mérés tanulást segítő funkciójának kihasználását, a tanulási folyamatba ágyazott, kis tétellel bíró diagnosztikus és formatív értékelést. Mindezen technológiai fejlesztések pontosabbá, objektívabbá, autentikusabbá, a diákok számára motiválóbbá tették a tesztelés folyamatát, illetve láthatóvá a tanulás hatékonyságát (Hattie 2009; Molnár–Csapó 2019c). A mérés-értékelés funkciójának és céljának újragondolási szükségességét erősítette fel a nagy tétellel bíró szummatív tesztekkel szemben a vészhelyzeti távolléti oktatás idején megfogalmazott kritika és az azzal párhuzamosan megjelenő igény az oktatás folyamatának, módszereinek, eszközeinek újragondolására.

Összességében megfogalmazható, hogy az elmúlt 30 évben a technológia – infrastruktúra, módszertan, eljárások – rapid fejlődése, ezzel összefüggésben a tanulásról és tudásról való tudásunk jelentős mértékű növekedése, illetve az értékesnek tartott tudás erőteljes változása következtében nemzetközi szinten szignifikáns átalakuláson ment át a mérés-értékelés célja, funkciója, tárgya, módszerei és eszközei, korábban nem lévő lehetőségek váltak elérhetővé. A hatékony tesztelés megvalósításának célját – annak teljesítése mellett – felváltotta a személyre szabott tanulás közvetlen támogatását megvalósító tesztelés. A nagy tétellel bíró szummatív tesztek korábbi dominanciája mellett egyértelműen megjelent a tanulási folyamatot közvetlenül támogató kis tétellel bíró formatív és diagnosztikus értékelés. A diszciplináris tudás kizárólagos mérése helyett előtérbe került a tudás alkalmazhatóságának, illetve a munkaerőpiaci igényeknek megfelelően, a gondolkodási képességeknek (l. 21. századi kulcskompetenciák) az innovatív mérése. A kizárólag válaszadatokon alapuló teljesítményértékelés mellett/helyett a kutatásokban egyértelműen megjelent a tesztelt személy feladatmegoldás során mutatott viselkedési mintázatainak elemzése és a személyre szabott vissza-

csatolás alkalmazása. Mindezen folyamatokat, változásokat ismét előtérbe helyezte és felgyorsította a mesterséges intelligencia széles körben való elérhetősége.

A mesterséges intelligencia szerepe a személyre szabott tanulás megvalósításában

Az iskola fő feladata, hogy diákjait felkészítse a 21. században való boldogulásra, felvértezze őket mindazon tudással és képességekkel, amelyek szükségesek a társadalomban, a munkaerőpiacon való eligazodáshoz és a sikeres életvezetéshez. Ezt a feladatot akkor tudja sikeresen betölteni, ha (1) a rendszer megfelelő tanulási kimeneteket definiál, (2) megfelelő eszközökkel rendelkezik annak diagnosztizálása céljából, hogy hol tartanak a diákok a tanulásban, illetve ha (3) a diákok tanulását, fejlesztését magasan képzett, tág módszertani repertoárral bíró, azaz magas minőségű pedagógusok támogatják.

A vészhelyzeti távolléti oktatás tapasztalatai, a technológia rapid fejlődése, beleértve a mesterséges intelligencia széles körű hozzáférhetővé válását, ismét felhívta a figyelmet a hagyományos keretek között zajló oktatás tárgyának, mikéntjének, módszereinek átgondolására. Számos kérdés, kétely és szélsőséges vélemény fogalmazódott meg a mesterséges intelligencia oktatási alkalmazása kapcsán, mely kételyek és kérdések jelentős részben az értékelés témakörén belül fogalmazódtak meg, holott a vonatkozó oktatási fejlesztések már a hatvanas évek végén elindultak (*Minsky–Papert 1968*), és ma már a mesterséges intelligencia számos eszközét használjuk a kutatásalapú oktatási és kis tétellel bíró értékelési rendszerekben (*Molnár–Csapó 2019a*).

Kis tétellel bíró rendszerek például a diákok fejlődését, tanulását közvetlenül támogató, azonnali, személyre szabott visszacsatolást nyújtó diagnosztikus rendszerek. A továbbiakban egy hazai példán – a Szegedi Tudományegyetem Oktatáselméleti Kutatócsoportjában és az MTA-SZTE Digitális Tanulási Technológia Kutatócsoportjában fejlesztett eDia-rendszeren keresztül – mutatom be a diagnosztikus rendszerek három fő funkcióját.

1. A rendszer tudományosan megalapozott, a diákok fejlődési sajátosságait figyelembe vevő elméleti keretrendszerre épülő, változatos, magas minőségű értékelési eszközöket – feladatokat, tesztek – biztosít mind a diákok, mind a pedagógusok részére (az eDia-rendszer vonatkozásában l. *Molnár et al. 2021; Csapó–Molnár 2019*). Egy kidolgozott rendszerben, a feladatok empirikus paraméterezése után a kiközvetített feladatok nehézségi szintje illeszkedik a diákok képesség-szintjéhez.
2. Különböző elemzéseken alapuló jelentéseket és azonnali visszajelzést nyújt a tanulás előrehaladásáról, a diákok adatai időben követhetőek és összehasonlíthatóak egymással. A rendszerhasználat után közvetlenül – a monitorozott területek kapcsán – a tanárok átfogó képet alkothatnak a diákok átlagos tudás- és képességszintjéről, hogy minél inkább megfelelő, személyre szabott oktatást nyújthassanak számukra (az eDia-rendszer vonatkozásában l. *Molnár–Csapó 2019a*).
3. A diagnosztikus rendszer harmadik fő funkciója már nemcsak a diagnózis felállítását támogatja, hanem a személyre szabott beavatkozás, fejlesztés megvalósítását is. Olyan online programokat tartalmaz (az eDia-rendszer esetében lásd az eLea adaptív fejlesztő programjai – <https://edia.hu/elea/>, illetve <http://edu.u-szeged.hu/dttk> vagy nemzetközi szinten a Khan Academy fejlesztései – <https://www.khanacademy.org/>), melyek akár tanórai, akár tanórán kívüli tevékenységekbe is integrálhatóak. A programok és a mesterséges intelligencia adta technológiák előnye, hogy nem fáradnak, ezért jó társa – nem helyettesítője! – lehet a pedagógusnak a személyre szabott oktatás megvalósításában.

A továbbiakban áttekintünk néhány olyan konkrét mesterséges intelligencia adta technológiai megoldást, amelyek már jelen vannak a hazai fejlesztésekben, így elérhetőek az eDia-platform keretein belül is, hasonlóan az azonos célú nemzetközi fejlesztésekhez, jelentősen csökkentve a pedagógusok terheit. Ugyanakkor lényeges megjegyezni, hogy az esetek egy részében alkalmazásukra még csak kutatási szinten került sor. A cél a tanulás folyamatának minél alaposabb megértése és a diákok között azonos évfolyamon meglévő, években mérhető különbségek mértékének csökkentése a személyre szabott oktatás alapját képező diagnosztikus értékelés lehetőségeinek folyamatos tágításával.

A mesterséges intelligencia adta minőségibb és érvényesebb értékelést megvalósító lehetőségek áttekintése

Az alábbiakban áttekintjük a mesterséges intelligencia adta minőségibb és érvényesebb értékelést megvalósító főbb lehetőségeket – mind a nagy adathalmazokkal dolgozó kutatás-fejlesztésen alapuló értékelési rendszerekben, mind a közvetlen osztálytermi értékelés esetén:

- *Automatikus, automatizált értékelés* – az emberi szubjektivitás, elfoglaltság befolyásoló hatásának kiiktatása. Az értékelés objektivitását, megbízhatóságát, pontosságát jelentősen növeli a válaszok automatikus értékelése. Például az automatizált szöveg/esszé értékelést lehetővé tevő NLP (natural language processing – természetes nyelvfeldolgozás) algoritmusok nemcsak az íráskészség sok szempontú értékelésének objektivitását és validitását javítják (Zhai–Ma 2023; Stadler et al. 2023), hanem személyre szabott visszajelzést adhatnak a szövegminőség javítása érdekében is.
- *A válaszadatok kiértékelését és elemzését követő gyors visszacsatolás.* Az értékelés hatékonyságát jelentős mértékben növeli a válaszadatok azonnali értékelése, elemzése, majd a gyors visszacsatolás. „Az MI alapú értékelési rendszer automatikusan értékeli a diákok válaszait és visszajelzést adhat a tanulás hatékonyságáról, annak javítása érdekében” (Duan–Tang–Liu 2022: 4).
- *A visszacsatolás gyorsasága mellett kulcsfontosságú a személyre szabott, könnyen értelmezhető visszacsatolás biztosítása.* A mérés-értékelés kulcsfontosságú eleme az eredmények jól érthető, könnyen értelmezhető, vizualizációs elemekkel ellátott formában történő visszacsatolása. A személyre szabott jelentések generálása során akár a képi, akár a szöveges megjelenítésben komoly támogatást adhat a mesterséges intelligencia, ezzel jelentősen növelhető a visszacsatolás tanulás-tanítási folyamatra gyakorolt hatékonysága.
- *Félautomatikus/automatikus feladatgenerálás* (Gierl–Haladyna 2013). A változatos, ugyanakkor megbízhatóan működő, minőségi feladatok nagy tömegű előállításának lehetősége a félautomatikus feladatgenerálás. Ebben az esetben a feladatok váza, szerkezete állandó. A feladaton belül előre definiált változókkal (szöveg, kép, hang, szám, multimédiás elem) dolgozunk. A rendszer random vagy egy adott algoritmusnak megfelelően generálja egy előre meghatározott halmazból a különböző értékű változókat a feladatokba. Miután a feladatok nehézsége jelentősen változhat a változók értékének változása következtében, ezért a feladatgenerálási funkciót inkább gyakorlásnál és fejlesztésnél, mint értékelésnél javasolt alkalmazni (Molnár–Makay–Ancsin 2018). A teljesen automatikus feladatgenerálás számos korláttal rendelkezik, ugyanis a szöveggenerátor-alapú mesterséges intelligencia-rendszerek nem ismerik az item- és tesztfejlesztés szabályait, nem vizsgálják a kidolgozott feladat értelmét, hanem valószínűségi alapon illesztik egymáshoz a szavakat, illetve azok kisebb egységeit. Pl. a ChatGPT-vel íratott feladat megfelelő alapja lehet a feladatírásnak, de a feladat lektorálása, felülvizsgálata nélkül automatikusan ma még nem alkalmazható megbízható adatok felvétele céljából, sőt a diákok számára adott téves, nem megfelelő stílusban generált visszacsatolás megakaszthatja a diák tanulási folyamatát.

- *Kontextuális adatok (log and process data) elemzése.* A diákok tanulási folyamatával kapcsolatos, a tanulás szempontjából lényeges lehetőségek a kontextuális adatok (pl. feladatmegoldással töltött idő, interakció milyensége, arckifejezéselemzés [Molnár–Lőrincz 2012] – a bevonódás, motiváció, frusztráció, szorongás szintjének megállapításához) mintázatkereső algoritmusokkal történő elemzése (process data elemzések). Kutatási eredmények alapján egyrészt az önbevalláson alapuló kérdésekre adott válaszok (kérdőív alapú mérés) sok esetben nem fedik a diákok valóságban mutatott viselkedését (l. Csányi–Molnár 2023), másrészt a kizárólag válaszadatokon alapuló képességszint-meghatározás téves következtetések levonását eredményezheti (részletesebben l. Molnár 2022). A mesterséges intelligencia e típusú elemzései teszik majd lehetővé a diákok érdeklődési körébe illeszkedő, motiváló, egyéni profiljára szabott (Gorgun–Bulut 2023) feladatok kiközvetítését (Liu–Baucham 2023).
- *Adaptív értékelés és tesztelés.* A tesztelt személy tudás- és képességszintjét fókuszba állító, ahhoz illeszkedő adaptív tesztek kiközvetítése, melyek a fix tesztekhez képest ugyanannyi idő alatt pontosabb mérést tesznek lehetővé (Csapó–Molnár–R. Tóth 2008; Magyar–Molnár 2013), vagy a mérési pontosságot megtartva jelentős mértékben – harmadával, felével csökkentik a tesztelési időt (Frey–Seitz 2009), miközben a diákok számára motiválóbbr tesztkörnyezetet biztosítanak.
- *Tanulásieredményeket előre jelző elemzések.* Nagy adathalmazok és gépi tanulási algoritmusok alkalmazásával különböző paraméterek ismeretében bizonyos valószínűség mellett előre jelezhető a tesztelt személy jövőbeli teljesítménye. Ezen elemzések a lemorzsolódás csökkentése, a tanulási sikeresség növelése szempontjából lényeges, fejlesztésre szoruló területek kijelölése kapcsán játszanak jelentős szerepet (Molnár–Csapó 2019b).
- *Automatizált beszédfelismerés.* Az NLP (natural language processing), azaz a természetes nyelvfeldolgozás funkciója a mérés-értékelés területén lehetőséget teremt például a nyelvi kompetenciák (pl. idegennyelv kapcsán kiejtés, kommunikációs készségek), kisgyermek mérése esetén a szóban adott válasz (pl. folytasd a számsort: egy, kettő, három, ...) automatizált kiértékelésére.

2023-ban a legszélesebb körben alkalmazott mesterségesintelligencia-program/modell, a ChatGPT oktatási/mérés-értékelésbeli alkalmazása során érdemes szem előtt tartani, hogy a kapott válaszok minősége, validitása, relevanciája a feltett kérdés minőségén múlik. Ha nem eléggé pontos a kérdés, a válasz sem lesz pontos. Ezenfelül a nagy nyelvi modellel rendelkező Chatbotok, mint a ChatGPT, nem vizsgálják a legenerált szöveg, válasz helyességét, valóságát. A pedagógus kiemelt feladata ezen eszköz alkalmazása során, hogy célzottan megtanítsa a diákokat arra, hogy minden esetben legyenek kritikusak és nézzenek utána a kapott eredményeknek.

A kutatók által fejlesztett értékelési rendszerek használatán túl Liu és Baucham (2023) javaslata alapján a pedagógus három megközelítéssel élhet, miközben a mesterséges intelligencia (vegyük példaként a ChatGPT-t) lehetőségeit integrálja az oktatás, illetve az oktatás folyamatán belül a mérés-értékelés folyamatába: 1. aktív, 2. passzív vagy 3. interaktív megközelítéssel.

Az aktív megközelítés alkalmazásával a pedagógusnak értenie kell az MI-technológia (pl. ChatGPT) oktatási funkcióihoz, ismernie kell korlátait, majd ezen jellemzőkre alapozva ki kell dolgoznia értékelési módszerét, a diákokat pedig arra ösztönözni, hogy legyenek kreatívok, kritikusak és eredetibbek, mint maga az MI-technológia.

A passzív megközelítés során a pedagógusok egy (valós) probléma megoldása kapcsán arra kéri a diákokat, hogy az MI-technológiától (pl. ChatGPT) gyűjtsenek információt/megoldási lehetőségeket, majd elemezzék azokat és érveljenek az általuk preferált megoldások mellett. A megoldás menetét és az eredményeket részletesen mutassák be társaiknak, kiemelve az adott MI használatának előnyeit és korlátait. Ez a megközelítés élvezetes tanulási folyamatot teremt a diákok számára, miközben fejleszti problémamegoldó képességüket, kritikus gondolkodásukat és kommunikációs

készségeiket is, ugyanakkor nehézséget jelenthet azon pedagógusok számára, akik egyáltalán nem ismerik az MI-technológia korlátait, fő jellemzőit.

Az interaktív megközelítés szerint a pedagógusok tudásmegosztás, közös gondolkodás céljából MI-technológia felhasználói csoportokat hoznak létre, ahova mind a pedagógusok, mind a diákok szabadon csatlakozhatnak. A csoport célja egy olyan értékelési módszer kidolgozása, ami mindkét fél szerint támogatja a tanulás és tanítás folyamatát. Ez a megközelítés egyértelműen megváltoztatja a hagyományos tanár-diák szerepeket és előtérbe helyezi a társas, együttműködésen alapuló tanulás fontosságát.

Konklúzió

Összességében megfogalmazható, hogy a tanítás és tanulás területén, beleértve a mérés-értékelést is, számos mesterségesintelligencia-alapú lehetőség érhető már el – leginkább kutatási szinten –, sőt alkalmazásuk kapcsán igen ígéretes eredményekkel rendelkezünk. Ugyanakkor technológia és mérés-értékelés kapcsán nagyon nagy a szakadék a kutatásban és a mindennapi iskolai gyakorlatban alkalmazott eszközök között. Lassú a lehetőségek gyakorlati adaptációja. Az MI – hasonlóan a többi új technológiához – oktatási alkalmazása annál hatékonyabb, minél inkább valós oktatási problémák megoldására törekszik (és nem csak a használat maga az oktatási cél), illetve minél jelentősebb az adott fejlesztő/értékelési eszköz pedagógiai, tanuláselméleteken alapuló beágyazottsága.

A technológiai, így a mesterséges intelligenciával kapcsolatos fejlesztések is nagyon gyorsan zajlanak, folyamatosan bővítve az MI oktatási problémák megoldásában történő használatának körét. Ugyanakkor a terület emberi tanulási folyamatokra gyakorolt hatása ismeretlen, azaz elementáris szükség van az MI emberi tanulásra gyakorolt hatásának alapos vizsgálatára. Sőt, minél inkább gyakori az MI tanulás során történő alkalmazása, annál inkább szükség van vonatkozó kutatásokra és a hatások elemzésére.

A mesterséges intelligencia adta lehetőségek (pl. gépi tanulás, nagy adathalmazok elemzése, mintázatkereső algoritmusok alkalmazása) mérés-értékelésbeli implementációjával a diákok számára személyre szabott, motiválóbb, interaktívabb tanulási, tesztelési környezet fejleszthető. Ennek hatására a sikeresen alkalmazható oktatási módszerek átalakulnak, hasonlóan, mint maga a pedagógusok tanulási-tanítási folyamatban betöltött szerepe. Miután tevékenységük egy részét támogatni tudják a technológiai eszközök, nagyobb hangsúlyt tudnak fektetni a minőségi oktatás alapját jelentő minél inkább személyre szabottabb oktatás megvalósítására.

A következő évtizedben valószínűleg jelentősen bővülni fog az MI-technológiák oktatási alkalmazása. Ezen eszközök a pedagógusok hatékony segítőivé tudnak válni. Ugyanakkor a különböző tanuláselméleteket ötvöző, hatékony tanulás ismérveit kereső kutatások, valamint a pandémia alatt tapasztalt távolléti oktatás hatását monitorozó kutatások eredményei szerint ezen technológiák nem fogják, nem tudják teljes mértékben kiváltani az emberi tanítást. Az iskola, a közösséggel, az osztálytársakkal való közös évek ugyanis jelentősen többről szólnak, mint ismeretek közvetítéséről.

Köszönetnyilvánítás

A tanulmány megírását az OTKA K135727 kutatási projekt, illetve az MTA Közoktatás-fejlesztési Kutatási Programja (KOZOKT2021-16) támogatta.

IRODALOM

ADESOPÉ, O. O. & RUD, A. G. (2019) Maximizing the affordances of contemporary technologies in education: Promises and possibilities. In: O. O. ADESOPÉ & A. G. RUD (eds) *Contemporary Technologies in Education*. Cham, Springer Nature. pp. 1–16. https://doi.org/10.1007/978-3-319-89680-9_1

- CSÁNYI R. & MOLNÁR GY. (2023) How do test-takers rate their effort? A comparative analysis of self-report and log file data. *Learning and Individual Differences*, Vol. 106. 10234. <https://doi.org/10.1016/j.lindif.2023.102340>
- CSAPÓ B. (2008) A magyar iskolarendszer adaptációs problémái: a tudás minősége. In: FAZEKAS K. (ed.) *Közoktatás, iskolai tudás és munkapiaci siker*. Budapest, MTA Közgazdaságtudományi Intézet. pp. 113–131.
- CSAPÓ B. & MOLNÁR GY. (2019) Online diagnostic assessment in support of personalized teaching and learning: The eDia system. *Frontiers in Psychology*, Vol. 10. p. 1522. <https://doi.org/10.3389/fpsyg.2019.01522>
- CSAPÓ B., MOLNÁR GY. & R. TÓTH K. (2008) A papíralapú tesztekől a számítógépes adaptív tesztelésig. *Iskolakultúra*, Vol. 18. Nos 3–4. pp. 3–16.
- DUAN, Y., TANG, H. & LIU, D. (2022) An AI-based assessment system for online learning: Design and implementation. *International Journal of Emerging Technologies in Learning*, Vol. 17. No. 1. pp. 4–13. <https://doi.org/10.3390/su141710551>
- FREY, A. & SEITZ, N. N. (2009) Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation*, Vol. 35. Nos 2–3. pp. 89–94. <https://doi.org/10.1016/j.stueduc.2009.10.007>
- GIERL, M. J. & HALADYNA, T. M. (2012, eds) *Automatic item generation: Theory and practice*. Routledge.
- GORGUN, G. & BULUT, O. (2023) Incorporating test-taking engagement into the item selection algorithm in low-stakes computerized adaptive tests. *Large-scale Assessments in Education*, Vol. 11. No. 17. <https://doi.org/10.1186/s40536-023-00177-5>
- GREENHOW, C., ROBELIA, B. & HUGHES, J. E. (2009) Learning, teaching, and scholarship in a digital age: Web 2.0 and classroom research: What path should we take now? *Educational Researcher*, Vol. 38. No. 4. pp. 246–259. <https://doi.org/10.3102/0013189X09336671>
- HATTIE, J. (2008) *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon, Routledge.
- JIMENEZ, K. (2023) 'This shouldn't be a surprise.' The education community shares mixed reactions to ChatGPT. *USA Today*. <https://www.usatoday.com/story/news/education/2023/01/30/chatgpt-going-banned-teachers-sound-alarm-new-ai-tech/11069593002/>
- LIU, Y. & BAUCHAM, M. (2023) AI Technology: Key to successful assessment. In: *Handbook of Research on Redesigning Teaching, Learning, and Assessment in the Digital Era*. IGI Global. pp. 304–325. <https://doi.org/10.4018/978-1-6684-8292-6.ch016>
- MAGYAR A. (2012) Számítógépes adaptív tesztelés. *Iskolakultúra*, Vol. 22. No. 6. pp. 52–60.
- MAGYAR A. & MOLNÁR GY. (2013) Adaptív és rögzített formátumú tesztek alkalmazásának összehasonlító hatékonyságvizsgálata. *Magyar Pedagógia*, Vol. 113. No. 3. pp. 181–193.
- MILLER, K. (2023) *Human writer or AI? Scholars build a detection tool*. *Stanford University Human-Centered Artificial Intelligence*. <https://hai.stanford.edu/news/human-writer-or-ai-scholars-build-detection-tool>
- MINSKY, M. & PAPER, S. (1968) *Linear separation and learning*. <https://dspace.mit.edu/bitstream/handle/1721.1/6170/AIM-167.pdf?sequence=2> [Letöltve: 2023. 08. 10.]
- MOLNÁR GY. (2011) Az információs-kommunikációs technológiák hatása a tanulásra és oktatásra. *Magyar Tudomány*, Vol. 172. No. 9. pp. 1038–1047.
- MOLNÁR GY. (2022) How to make different thinking profiles visible through technology: The potential for log file analysis and learning analytics. In: M. VIRVOU, G. A. TSIHRINTZIS, L. H. TSOUKALAS & L. C. JAIN (eds) *Advances in Artificial Intelligence-based Technologies*. Cham, Springer. pp. 125–145. https://doi.org/10.1007/978-3-030-80571-5_9
- MOLNÁR GY. & CSAPÓ B. (2019a) A diagnosztikus mérési rendszer technológiai keretei. Az eDia online platform. *Iskolakultúra*, Vol. 29. Nos 4–5. pp. 16–32. <https://doi.org/10.14232/ISK-KULT.2019.4-5.16>

- MOLNÁR GY. & CSAPÓ B. (2019b) A felsőoktatási tanulmányi alkalmasság értékelésére kidolgozott rendszer a Szegedi Tudományegyetemen: elméleti keretek és mérési eredmények. *Educatio*, Vol. 28. No. 4. pp. 705–717. <https://doi.org/10.1556/2063.28.2019.4.4>
- MOLNÁR GY. & CSAPÓ B. (2019c) Making the psychological dimension of learning visible: Using technology-based assessment to monitor students' cognitive development. *Frontiers in Psychology*, Vol. 10. p. 1368. <https://doi.org/10.3389/fpsyg.2019.01368>
- MOLNÁR GY. & CSAPÓ B. (2023) Report on smart education in Hungary. In: R. ZHUANG et al. (eds) *Smart Education in China and Central & Eastern European Countries. The white paper on smart education in China and Central & Eastern Europe countries*. Springer, Singapore. pp. 155–178. https://doi.org/10.1007/978-981-19-7319-2_7
- MOLNÁR GY. & LŐRINCZ A. (2012) Innovative assessment technologies: Comparing 'face-to-face' and game-based development of thinking skills in classroom settings. In: *International proceedings of economics development and education innovation*. Vol. 37. pp. 150–154.
- MOLNÁR GY., MAKAY G. & ANCSIN G. (2018) *Feladat- és tesztszerkesztés az eDia-rendszerben*. Szeged, Oktatásméleleti Kutatócsoport.
- MOLNÁR GY., PÁSZTOR A., KISS R. & CSAPÓ B. (2021) Az eDia online diagnosztikus értékelő rendszer: a személyre szóló fejlesztés alapvető eszköze. *Új Pedagógiai Szemle*, Vol. 71. Nos. 9–10. pp. 42–53.
- MOLNÁR GY., TURCSÁNYI-SZABÓ M. & KÁRPÁRI A. (2019) Az interaktív tanulási környezetektől a módszertani megújulásra át a kreatív önkifejezésig. *Új Pedagógiai Szemle*, Vol. 69. Nos 11–12. pp. 53–70.
- MOLNÁR GY., TURCSÁNYI-SZABÓ M. & KÁRPÁRI A. (2020) Digitális forradalom az oktatásban – perspektívák és dilemmák. *Magyar Tudomány*, Vol. 181. No. 1. pp. 56–67. <https://doi.org/10.1556/2065.181.2020.1.6>
- PRESSEY, S. L. (1927) A machine for automatic teaching of drill material. *School & Society*, Vol. 25. pp. 549–552.
- REDECKER, C. & JOHANNESSEN, Ø. (2013) Changing assessment – Towards a new assessment paradigm using ICT. *European Journal of Education*, Vol. 48. No. 1. pp. 79–96.
- SINGH, H. & MIAH, S. J. (2020) Smart education literature: A theoretical analysis. *Education and Information Technologies*, Vol. 25. No. 4. pp. 3299–3328. <https://doi.org/10.1007/s10639-020-10116-4>
- STADLER, M., HÖRRER A, FISCHER, M. R. & SAILER, M. (2023) *Utilizing large language models for coding open-ended responses in psychological assessment: An empirical investigation*. Preprint.
- SVRLUGA, S. (2023) Was that essay written by AI? A student made an app that might tell you. *The Washington Post*. <https://www.washingtonpost.com/education/2023/01/12/gptzero-chatgpt-detector-ai/>
- The Associated Press (2023) *Amid ChatGPT outcry, some teachers are inviting AI to class*. <https://nowthisnews.com/news/amid-chatgpt-outcry-some-teachers-are-inviting-ai-to-class> [Letöltve: 2023. 08. 09.]
- ZHAI, N. & MA, X. (2023) The effectiveness of automated writing evaluation on writing quality: A meta-analysis. *Journal of Educational Computing Research*, Vol. 61. No. 4. pp. 875–900. <https://doi.org/10.1177/07356331221127300>

A cikk a Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>) feltételei szerint publikált Open Access közlemény, melynek szellemében a cikk bármilyen médiumban szabadon felhasználható, megosztható és újraközölhető, feltéve, hogy az eredeti szerző és a közlés helye, illetve a CC License linkje és az esetlegesen végrehajtott módosítások feltüntetésre kerüljenek. (SID_1)