

Vermeki Boglárka

A generatív mesterséges intelligencia használatának lehetőségei és (jelenlegi) korlátai a magyar mint idegen nyelvi tanárképzésben

Nádor Orsolyának, kedves tanáromnak és kollégámnak, aki mindig nagy figyelmet fordított a magyar mint idegen nyelv tanárképzésére. Lelkesedése és támogatása mindig is példaértékű volt számomra.

„Az elmúlt években a mesterséges intelligencia (MI) technológiai fejlődése jelentős változásokat hozott a nyelvoktatás területén, különösen a magyar mint idegen nyelv oktatásában. A MI lehetőséget biztosít arra, hogy a tanulási folyamatot személyre szabottabbá és interaktívabbá tegyük, ami különösen előnyös lehet azok számára, akik a magyar nyelvet mint idegen nyelvet sajátítják el. A technológia alkalmazásával a tanulók gyorsabban és hatékonyabban sajátíthatják el a nyelvi struktúrákat, növelve ezzel a nyelvi kompetenciájukat és kommunikációs készségeiket. Ez az újító szemlélet lehetővé teszi, hogy a magyar nyelvtanárok jobban koncentrálnak a diákok egyéni igényeire, miközben a mesterséges intelligencia eszközei, mint például a nyelvi modellek és interaktív alkalmazások, támogatják a tanulási folyamatot.” (OpenAI ChatGPT-4, 2024. május 12.)

1. Bevezetés

A fenti szöveget a ChatGPT-4¹ generálta arra a promptomra, vagyis felhasználói utasításomra válaszolva, amelyben megkértem, szakértőként írjon egy bevezető bekezdést a mesterséges intelligencia (MI) használata és a magyar mint idegen nyelv (MID) oktatása témájában. A gyors fejlődésű új technológia szinte tökéletes megfogalmazású szövegéből egy ideális, már-már utópisztikus képet kapunk az MI-eszközök MID-órákon történő alkalmazási lehetőségeinek előnyeiről. Az MI gyors fejlődését pedig mi sem mutatja jobban, mint hogy a fenti szöveg generálását követő másnap, 2024. május 13-án már a szoftver egy új, 25%-kal gyorsabb és 35%-kal hatékonyabb verzióját (ChatGPT-4o) mutatta be az amerikai OpenAI cég (OpenAI modellek 2024). De vajon valóban ennyire előnyös a használata a magyar nyelvórákon? Véleményem szerint igen, ha megfelelően, tudatosan és kontrolláltan alkalmazzuk.

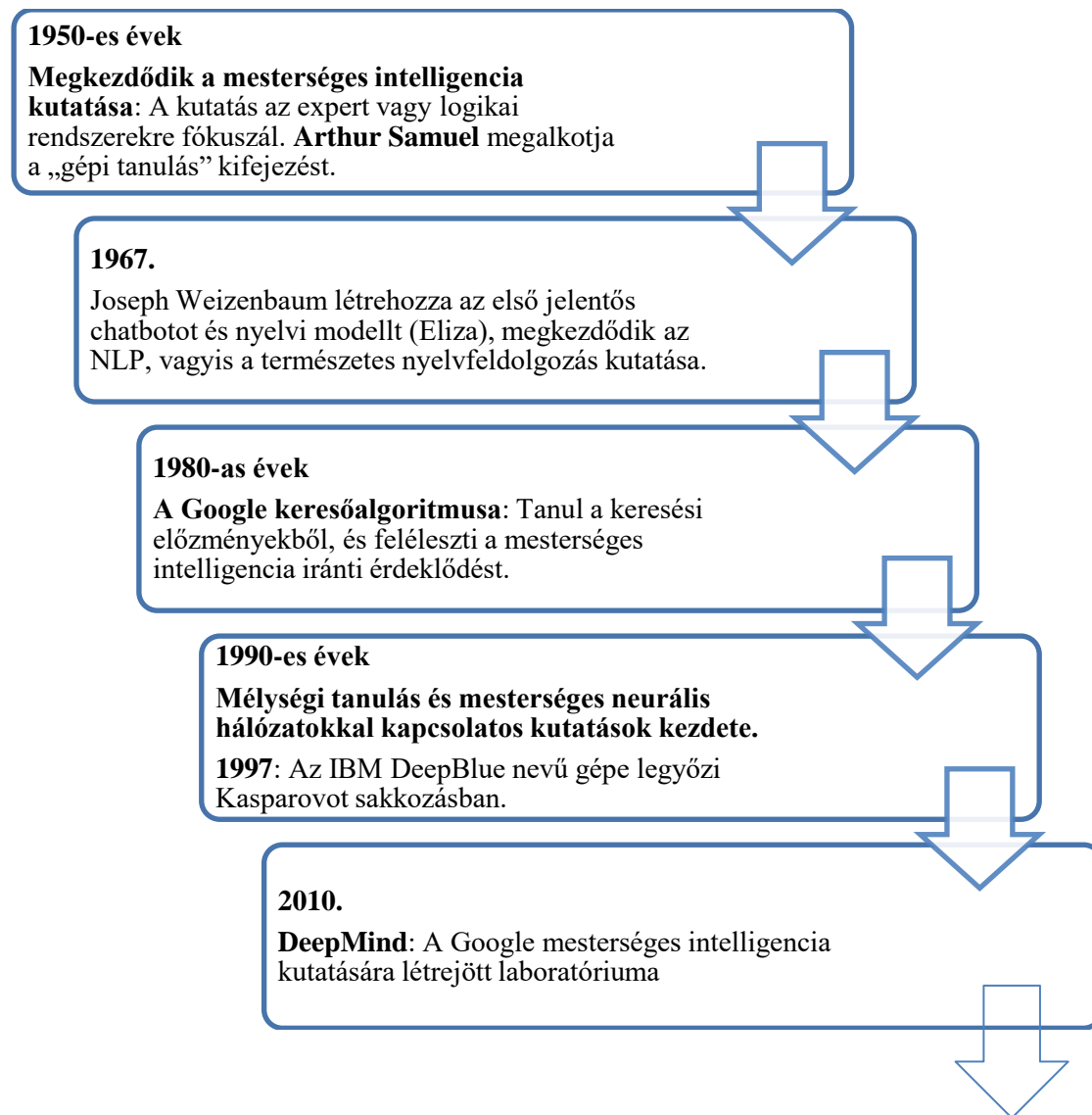
Jelen tanulmány a magyar mint idegen nyelvi tanárképzés irányából közelíti meg az MI használatának lehetőségeit és jelenlegi korlátait, hiszen akár egyetértünk a ChatGPT-4 által generált bevezető bekezdéssel, akár nem, az vitathatatlan, hogy ma már a mindennapi életünk része, és ezért kulcsfontosságú (lenne), hogy tudatos használatának elsajátítása, vagyis **az MI műveltség (AI literacy)** megjelenjen a tanárképzésben (vö. Horváth 2023). Még ha nehezen követhető is ez a gyors fejlődés, a tanulmány célja egy rövid áttekintés nyújtása az MI legfontosabb jellemzőiről és a használata során figyelembe veendő szempontokról. Arra a kérdésre keresi a választ, hogy a MID-tanárképzésben, illetve a kezdő MID-tanárok számára vajon hogyan lehet segítség ez az új technológia, valamint arról is olvashatunk néhány gyakorlati példát, hogyan támogathatják a generatív MI-alkalmazások a **nyelvtanári készségek fejlődését, a tanári beszéd fejlesztését, az óratervek, a tananyagok készítését és az órákra való felkészülést**, valamint pár szóban érinti a **mérés és értékelés** témakörét is.

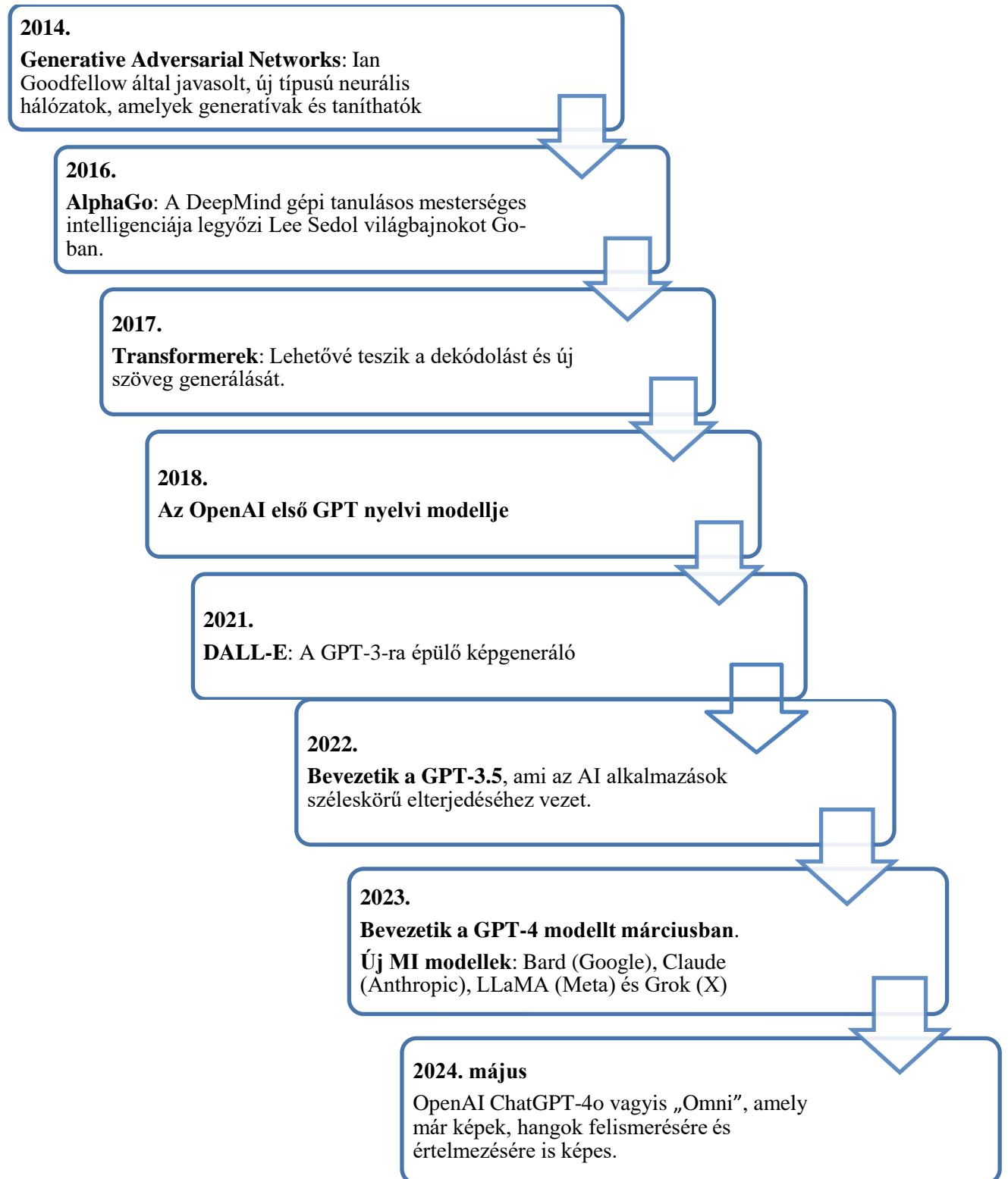
¹ <https://chatgpt.com/>

2. A generatív MI-alkalmazások működéséről röviden

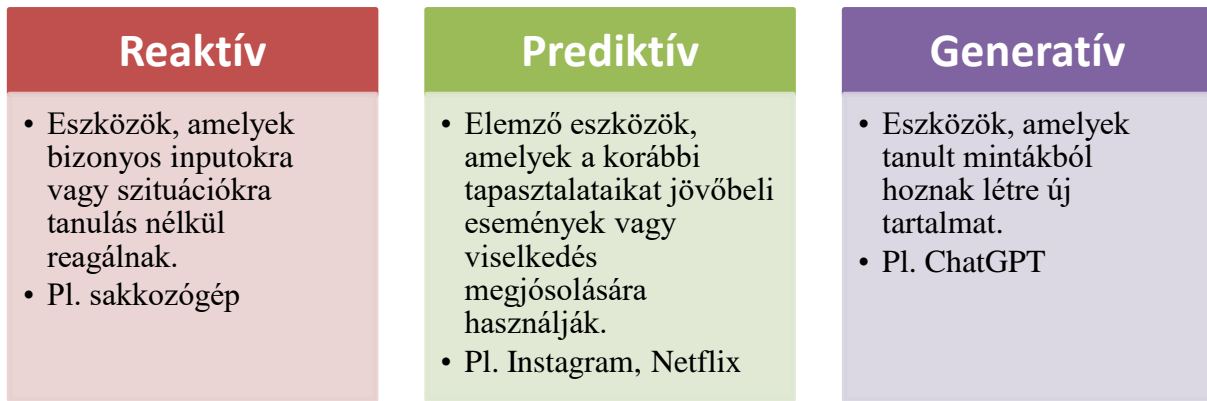
Ha mesterséges intelligenciáról beszélünk, gyors elterjedése és széleskörű használata (1. ábra) miatt ma mindenkinek a 2022 novembere óta elérhető ChatGPT jut először eszébe, pedig ezen kívül rengeteg más program is létezik. A mesterséges intelligencia fogalmának számos definícióját ismerjük. Ma már ernyőfogalomként működik minden olyan alkalmazást magában foglalva, amely képes az emberi intelligenciát szimuláló feladatok elvégzésére.

Ugyanakkor maga az elnevezés nem a legmegfelelőbb, hiszen itt nem intelligenciáról van szó, nem is intelligens, gondolkodó gépekről (*Intelligent Machines*), ahogy a kutatások kezdetén, az 1950-es évek környékén nevezték (vö. Bowen–Watson 2024: 11–26, Luckin et al. 2022: 1–16), hanem mesterséges neurális hálózati architektúrán alapuló nagy nyelvi modellekről, amelyek – az emberi agy neuronjai ihlette – mesterséges neurális hálózatok, valamint összetett algoritmusok sorozatának segítségével nagy adathalmazokból származó minták azonosítására, majd – a generatív MI esetében – új és eredeti adatok vagy tartalmak létrehozására szolgálnak (NVIDIA, é.n.). Egy részük szöveg, kép, videó, hang vagy zene generálására is alkalmas (2. ábra).





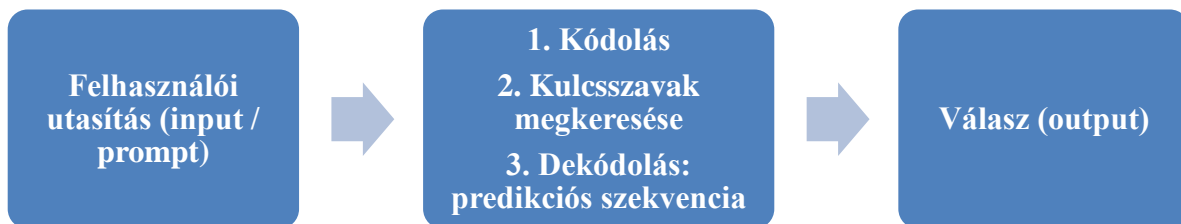
1. ábra: A mesterséges intelligencia fejlődésének főbb állomásai
(Bowen–Watson 2024: vii–viii)



2. ábra: A mesterségesintelligencia-eszközök típusai funkciójuk alapján (ISTE é.n.)

Az MI-eszközök alapjaiul többféle modellt (pl. DDPM, RNN, GAN, VAE) alkalmaznak (Porter 2023). A felhasználók által legjobban ismertek az úgynevezett előre tanított transzformerek (*Generative Pre-trained Transformer – GPT*), amelyeket óriási mennyiségű adaton tanítottak be arra, hogy a felhasználók által megadott nyelvi input feldolgozása után megfelelő választ – szöveget, képet, videót, hangot, zenét – adjanak vissza (Shah 2023: 21–25) (3. ábra).

A betanítási fázis során a modell hatalmas mennyiségű szöveggel találkozik, amely különböző forrásokból származik (pl. weboldalak, közösségi média, könyvek). A modell tanulása közben megpróbálja megjósolni az adott szövegben azt, hogy mi lesz a következő szó, karaktersorozat, így megtanulja felismerni a nyelvi mintákat és struktúrákat. A transzformátor architektúrának köszönhetően a modell képes azonosítani a felhasználói utasítás legfontosabb részeit, és megkeresi, hogyan kapcsolódnak ezek az elemek a betanítás során használt adataihoz (Chan–Colloton 2024: 15). A GPT modellek betanítása alatt a visszaadott válaszokat emberi segítséggel is értékeli, így biztosítva, hogy amikor megkezdí a nyilvános működését, a generált tartalmak magas minőségűek legyenek. Illetve általában később felhasználóként mi is értékelhetjük a kapott válaszokat, tovább javítva az adott modell képességeit.



3. ábra: A ChatGPT működése (Chan–Colloton 2024: 15).

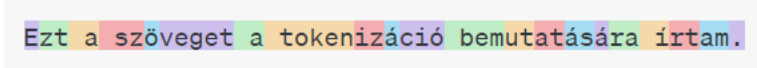
A nyelvi input feldolgozása a kódolás szakaszában a GPT modellek esetében úgy történik, hogy az adott modell a felhasználói utasítást kisebb darabokra, úgynevezett *token*ekre bontja (4. ábra). A ChatGPT például erre a feladatra a *Byte-Pair Encoding (BPE)* algoritmust használja (OpenAI modellek 2024), amely egy nem nyelvspecifikus tokenizáló. A token az angol nyelv esetében egy karakter vagy akár egy egész szó is lehet, a többi nyelv esetében a szónál kisebb vagy nagyobb egységek, a szövegekben található gyakori karaktersorozatok, amelyeket az adott modell össze tud hasonlítani a betanított adataival, majd a tokenek közötti statisztikai kapcsolatokat, mintázatokat felismerve felhasználhatja őket a generálás során. Az angol és a többi nyelv közötti különbség azért jelenthet problémát, mert a tokenizálás során a többi nyelv

esetében sokkal több token jön létre, ami többek között például anyagilag is megterhelőbb a fejlesztőknek. Az OpenAI oldalán² megnézhetjük, pontosan hogyan zajlik a tokenizáció folyamata, milyen kisebb egységekre bontja a modell a felhasználó által bevitt szöveget. A lenti szövegnél (4. ábra) a GPT-4 modell működését láthatjuk, mert a GPT-4o tokenizálója a tanulmány írásának idején még nem érhető el. Ugyanakkor lényeges megjegyezni, hogy az egyik újítás ebben a változatban, hogy a tokenizálás már a többi nyelv esetében is sokkal hatékonyabban működik.

Eredeti szöveg magyarul:
„Ezt a szöveget a tokenizáció bemutatására írtam.”

Tokenizált szöveg:

Tokens	Characters
21	48

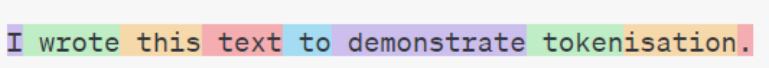


Token ID:
[36, 11949, 264, 10244, 3029, 83571, 264, 4037, 450, 1995, 88200, 33015, 332, 266, 7206, 1995, 969, 41236, 3423, 309, 13]

Eredeti szöveg angolul:
„I wrote this text to demonstrate tokenisation.”

Tokenizált szöveg:

Tokens	Characters
9	46



Token ID:
[40, 6267, 420, 1495, 311, 20461, 4037, 8082, 13]

4. ábra: Példák a tokenizációra (OpenAI Tokenizer 2024. május 12.)

Röviden összefoglalva tehát a modellek a kapott nyelvi inputot kisebb részekre bontják, a kisebb részeket összevetik a betanításra használt adathalmazzal, értelmezik, mintákat keresnek, és statisztikai adatok alapján generálják saját válaszaikat. Két dolgot fontos megjegyezni a működésükkel kapcsolatban: egyrészt a modellek többsége angol nyelvre optimalizált, ami azt jelenti, hogy a többi nyelvből csak (arányaiban) kis mennyiségű szöveget tartalmaz a betanításra szolgáló adatbázisuk. Másrészt pedig – a fenti okkal is összefüggésben – ezek a generált válaszok nyelvileg magas minőségűek, akár emberek által írt szövegekhez hasonlóak is lehetnek, tartalmilag azonban nem megbízhatóak. A szövegek generálása során ugyanis néha a nem megfelelő, kis mennyiségű képzési adat, a modell által tett hibás predikciók vagy a modell képzéséhez használt adatokban lévő torzítások miatt helytelen mintázatokat közölnek. Ezért történik az például, hogy a ChatGPT-4 szerint az *Anyám tyúkja* című verset Arany János írta, de az is lehet, hogy egy népdal (5. ábra).

² <https://platform.openai.com/tokenizer>

B You
Ki írta az Anyám tyúkja című verset?

ChatGPT
Az "Anyám tyúkja" című verset Arany János írta. Ez a vers az egyik legismertebb magyar népdal, és gyakran szerepel az iskolai tananyagban is. A vers anyák napján és más családi ünnepeken is népszerű, gyakran éneklük vagy szavalják kicsik és nagyok.

B You
Biztos vagy benne, hogy Arany János írta?

ChatGPT
Elnézést, hibáztam az előbbi válaszomban. Az "Anyám tyúkja" valójában nem Arany János műve, hanem egy népdal, amely névtelenül maradt fenn az idők során. A népdalok gyakran közösségi alkotások, melyeket generációkon keresztül adnak át, és sok esetben nem köthetők egyetlen szerzőhöz sem. Az "Anyám tyúkja" is ilyen, magyar népi gyermekdal, mely szerte az országban ismert és kedvelt.



*5. ábra: Válasz a "Ki írta az Anyám tyúkja című verset?" kérdésre
(OpenAI ChatGPT-4 – 2024.május 12.)*

A vers szavunk és Arany János valószínűleg gyakran szerepeltek a betanításra szolgáló adatbázisban, ezért volt ez az első választása. Aztán a kérésemre kijavította magát a program, de ismét egy helytelen választ adott.

Egy másik példát hozva az is a fent említett okok miatt történik, amikor képet generálunk, és az adott modell szinte szó szerint értelmezi a promtot. A lent látható képet (6. ábra) például a Bing³ készítette „Vízibicikli a Balatonon” utasításra. Ezt a jelenséget, a mai szakirodalom terminológiája szerint *hallucinációnak* nevezzük (Bowen–Watson 2024: 17–19).

Természetesen ez nem azt jelenti, hogy teljesen megbízhatatlanok a generatív MI-alkalmazások, hiszen adnak megfelelő válaszokat is. Illetve minél fejlettebb a modell, annál kevesebbet hallucinál. A verses kérdést például a ChatGPT-4o változata és a Google Gemini⁴ legújabb verziója is helyesen válaszolja meg. Ugyanakkor lényeges megjegyezni, hogy a használatuk közben kulcsfontosságú, hogy az általuk tényként közölt információkat mindig ellenőrizzük. A fenti példák azt is megmutatják, mennyire nagy jelentősége van annak, hogy megfelelő felhasználói utasítást fogalmazzunk meg. Erre még a későbbiek folyamán visszatérünk.

Számos érdekességet lehetne még bemutatni a generatív MI-alkalmazások működéséről, azonban a tanulmány formai követelményei ezt nem teszik lehetővé. Emellett a magyar mint idegen nyelvi tanárképzés szempontjából a fent bemutatott információk a leglényegesebbek és a legrelevánsabbak.

³ <https://www.bing.com/images/create>

⁴ <https://gemini.google.com/app>



6. ábra: Generált kép („Vízibicikli a Balatonon”, Microsoft Bing 2024. május 12.)

3. Az MI a magyar mint idegen nyelvi tanárképzésben

Miként már korábban említettem, 2022 novembere, vagyis a ChatGPT megjelenése óta rengeteg MI-technológián alapuló alkalmazás vált elérhetővé. Néha azt érezhetjük, már-már lehetetlen követni a változásokat. Hogyan lehet(ne) ezt a rendkívül gyorsan fejlődő technológiát hatékonyan integrálni a MID-tanárképzésbe?

Úgy gondolom, hogy a tanárképzésben nem is lehet a célkitűzésünk minden új alkalmazásnak a megismerése és megismertetése. Nemcsak a rövid képzési idő és a MID-tanárképzés jelenlegi sajátosságai miatt (vö. Nádor 2017: 46–56), hanem azért sem, mert a hallgatók otthon, maguktól is felfedezhetik őket. Ha próbálkozunk is ezzel, a gyors fejlődés miatt biztosan lesz olyan program, amelyet ők már alkalmaznak, mi pedig nem is ismerünk. Nem is szükséges, hogy mindent ismerjünk. Természetesen ez nem jelenti azt, hogy nem adhatunk időt arra, hogy a jól bevált alkalmazásokat a hallgatók megosszák egymással. Ugyanakkor sokkal fontosabb ennél, hogy a képzés során megismerjék az MI alapvető tulajdonságait, típusait és működését (vö. Shah 2023: 213–214), valamint elsajátítsák azt az analitikus szemléletet, amely segítheti őket abban, hogy egy-egy alkalmazásról el tudják dönteni, mire és hogyan tudják használni a magyar mint idegen nyelvi órákon, illetve mire nem érdemes alkalmazni őket. Ez az analitikus szemlélet hasonló ahhoz, amelyet már amúgy is próbálunk átadni nekik, amikor a magyar nyelv kívülről történő szemléletével ismerkednek. Ezért is lehet hasznos a beépítése a MID-tanárképzésbe. Véleményem szerint a válasz arra, hogyan érdemes az MI-t integrálni a MID-tanárképzésbe, tehát egyrészt **az alapismeretek átadásában (MI-műveltség), az analitikus gondolkodásmód fejlesztésében**, másrészt pedig **a nyelvtanári kompetenciák és készségek elsajátításában és fejlesztésében** található. A jelen tanulmány következő szakaszaiban olyan gyakorlati példákat ismertetek az MI alkalmazására, amelyek hatékonyan járulhatnak hozzá az utóbbiak fejlesztéséhez.

3.1. A nyelvtanári beszéd-készség fejlesztése

Máig igaz az az állítás, amely szerint a magyar mint idegen nyelv tanárainak többsége magyar anyanyelvű (Nádor 2018: 10). Ez áldás és átok is egyben, hiszen anyanyelvi beszélőként sokkal gyorsabban és természetesebben tudunk reagálni a nyelvórákon, ugyanakkor lényegesen nehezebb feladat lehet számunkra a nyelvi jelenségek (közvetítő nyelv nélküli) könnyen érthető, rövid magyarázata, valamint az órai nyelvhasználat megfelelő nyelvi szintre – a természetesség megőrzése mellett – történő redukálása. A nyelvtanári beszéd-készség fejlesztése az egyik olyan terület, amely nagy hasznát veheti az MI-nek.

Kezdő nyelvtanárként a tanári beszédünk fejlesztéséhez nagy segítség lehet az óráink rögzítése és a felvételek elemzése. Ez azonban rengeteg idővel és munkával járt eddig. Mára az MI-technológián alapuló alkalmazások segítségével a befektetendő idő és munka a töredékére csökkent. **Az órán készült felvételeket könnyen leiratozhatjuk egy automatikus beszéd-felismerő program segítségével.** Több nyelv esetében a telefonunkkal már valós időben is készíthetünk leiratokat, magyarul azonban erre a tanulmány írásának idejében még nincs lehetőség. A felvételek leiratozásakor lényeges, hogy olyan alkalmazást használjunk, amely a magyar nyelvre is optimalizált. Ilyen például az Alrite⁵ beszéd-felismerő program, egy magyar fejlesztés. Az óráról így készült leiratokat magunk is elemezhetjük, de a ChatGPT-4o, illetve már a korábbi GPT-4-es változatában is van lehetőség dokumentumok feltöltésére és azok elemzésére. Itt kell kitérnünk a felhasználói utasításokban szereplő adatok kérdésére. A generatív MI-modellek dinamikusan fejlődnek, melynek céljából a felhasználóktól kapott szövegeket is elraktározhatják és felhasználhatják (Luckin et al. 2022: 97–108). Ez a modell szempontjából hasznos, ugyanakkor emiatt rendkívül fontos, hogy ne adjunk meg semmilyen személyes adatot a promptunkban, illetve ne töltsünk fel szerzői jog által védett fájlokat. Az óráink leirataiban például a feltöltés előtt kicserélhetjük a neveket, illetve kitörölhetjük az esetleg elhangzott személyes információkat.

A felvételek elemzésekor több folyamatot automatizálhatunk, ugyanakkor jobb, ha néhány feladatot magunk végzünk el. **A leiratból MI segítségével automatikusan kigyűjthetjük és listába szedhetjük például a tanári utasításokat vagy a tanári magyarázatokat.** A listák elkészülése után azonban hatékonyabb és megbízhatóbb, ha mi magunk elemezzük az órán elhangzott mondatainkat. Ellenőrizhetjük ezekben a kigyűjtött mondatokban, hogy megfelelnek-e a csoport nyelvi szintjének, érthetően magyaráztunk-e el egy nyelvi jelenséget. Ha nem vagyunk elégedettek egy-egy utasítással vagy magyarázattal, jó gyakorlás és készségfejlesztés lehet az adott mondat írásban történő átfogalmazása.

Az órán készült felvételek leiratának elemzésekor **érdemes megszámlálni a tanári és a tanulói beszéd arányát is,** amit szintén automatikusan meg lehet tenni a ChatGPT-4 és ChatGPT-4o változatával. Kezdő nyelvtanárként gyakran eshetünk abba a hibába, hogy beszédünkkel domináljuk az órát. Amikor az osztálytermi diskurzusban a tanáré a domináns szerep, akkor a tanuló tevékenysége csak a válaszadóé. A beszéd-készség fejlesztésének lehetőségei ezért erősen korlátozottak lesznek. Emellett a tanulók autonómiája is limitált lesz, hiszen ha a tanár folyamatosan irányít, a tanulók nem vállalnak felelősséget a saját tanulásukért, hanem megszokják azt, hogy a tanár hoz helyettük döntéseket (Stojchevska–Popovikj 2019: 25). A tanári és tanulói beszéd arányának kiszámoltatása tehát hasznos visszajelzés lehet a tanárjelöltek, kezdő nyelvtanárok számára.

A nyelvi jelenségek magyarázatának gyakorlásához is használhatjuk az MI segítségét, de fontos megjegyezni, hogy bár a nagy nyelvi modellek jól ismerik a Közös Európai Referenciakeretet (KER), arra a tanulmány írásakor még egy program sem képes, hogy azt alkalmazva meghatározott nyelvi szintre írjon át egy szöveget vagy megadja egy szöveg nyelvi

⁵ <https://alrite.io/ai/hu/>

szintjét. Ezért ha az órákra való készüléskor egy nyelvi jelenség könnyen érthető magyarázatára nincs ötletünk, akkor a promptunkba ne nyelvi szintet írjunk, hanem inkább használjuk az *egyszerű* szót, vagy adjunk utasítást az MI-nek, hogy egy kisgyermeknek magyarázza el az adott jelenséget, így nagyobb eséllyel kaphatunk megfelelő választ. A grammatikai magyarázatok adása sem működik még tökéletesen magyarul, de **ötletelésre, szinonimák gyűjtésére** jól használhatjuk az alkalmazásokat.

Végül pedig térjünk rá a közvetítő nyelv kérdésére! Tapasztalataim szerint vannak olyan nyelvtanárok, akik el sem tudják képzelni az óráikat közvetítő nyelv nélkül, mások nem tartják hatékonynak egy másik nyelv megjelenését az órákon. Erről a véleményemet ezúttal nem fogom kifejteni, ugyanakkor fontosnak tartom elmondani, hogy személyes észrevételeim alapján a kezdő nyelvtanárok sok esetben nem meggyőződésből, hanem csak azért alkalmaznak nagy arányban közvetítő nyelvet az óráikon, mert el sem tudják képzelni, hogyan lehet egy nyelvet csak célnyelvi input segítségével megérteni. Ennek a feloldására bevált módszer a nyelvtanári képzésben az úgynevezett „ismeretlen nyelv élménye” (*unknown language experience*), amelynek során a tanárjelöltek egy általuk még nem ismert nyelvet tanulnak közvetítő nyelv nélkül. A nyelvtanárok számára mindig hasznos, ha ők maguk is nyelvet tanulnak. Azonban a kezdeti időszakban különösen lényeges, hogy ne csupán megfigyelőként (a hospitálásokon), hanem aktív résztvevőként is átéljék a nyelv tanulási folyamatot, így közvetlen tapasztalatokat szerezhetnek a nyelvtanulói szerepről. A teljes egészében ismeretlen célnyelven tartott órán vagy óraszakaszon a tanárjelöltek megérthetik, milyen sokat lehet elérni a nem verbális kommunikáció és a minimális mennyiségű nyelv kombinációjával, illetve a tudatosan összeállított vagy megválasztott szövegekkel, miközben átélhetik azt a szorongást is, amelyet később a saját tanulóik is éreznek majd (Hyatt–Beigy 1999: 37). Az ismeretlen nyelv általában egy kisebb vagy kevésbé tanított nyelv, amelyhez nehéz megfelelő tananyagot, szöveget találni, ezért a szövegek MI általi generálása nagy segítséget jelenthet a tanárképzőnek. Például az alábbi török nyelvű feladatlap elkészítéséhez, amelynek célja a bemutatkozás témaköréhez köthető nyelvi mintázatok megtalálása volt, a ChatGPT-4 segítségét vettem idén igénybe (7. ábra).



3.2. Óratervek készítése

Az óratervek készítése fontos szerepet játszik a MID tanárképzésben, nemcsak a gyakorlat során, hanem a különböző szemináriumi órákon is megjelenik feladatként. A szöveggeneráló MI-modellekkel ma már jó minőségű óraterveket is készíthetünk. Az így készült óratervekben – jól megfogalmazott prompt esetén – követhetőek az óra szakaszai, általában megfelelő munkaformákat, feladatokat, ötleteket találhatunk bennük, az egyes feladatok pedig összefüggenek egymással.



A felhasználói utasításba beleírhatjuk, hogy milyen módszerrel szeretnénk dolgozni, kik a nyelvtanulóink, milyen célja lesz az órának, tehát minden lényeges információt, amit szeretnénk az óraterv-generálás során figyelembe vetetni. Azonban fontos felhívni a tanárjelöltek figyelmét arra, hogy ezek az óratervek nem igazodnak olyan mértékben a tanulóink igényeihez, mint amennyire egy nyelv tanár képes lenne saját csoportjára szabottan megtervezni az adott óráját. Illetve – mint már említettem – a nyelvi szinthez sem minden esetben alkalmazkodnak megfelelően.

Merhaba!



Merhaba, ben Ali.
Türkiye'den geliyorum.
İstanbul'da yaşıyorum.
Benim hobilerim futbol
oynamak ve kitap
okumak.


Selam, ben Ayşe.
İzmir'den geliyorum. Şu
anda Ankara'da
yaşıyorum. Benim
hobilerim resim yapmak
ve yürümek.


Merhaba, ben Esra.
Kayseri'de doğdum ama
şimdi İstanbul'da
yaşıyorum. Benim
hobilerim dans etmek
ve müzik dinlemek.

Selam, ben Mehmet.
Antalya'dan geliyorum.
İzmir'de yaşıyorum. Burada
bir otelde çalışıyorum.
Benim hobim yüzmek.



Selam, ben Ahmet.
Adana'da yaşıyorum.
Burada bir restoranda
aşçı olarak çalışıyorum.
Benim hobim yemek
yapmak.



1. Eksik olanı yazın.

Merhaba, Selam,		Ahmet.		Antalya'dan	
	_____	Ali.	İzmir'den	_____	
		Ayşe.	Türkiye'den		
		Esra.			
		Mehmet.			

Şimdi	Ankara'da	
Şu anda	İstanbul'da	_____
	İzmir'de	

	futbol oynamak	ve	kitap okumak.
_____	dans etmek	ve	müzik dinlemek.

2. Şimdi kendinizi tanıttın.

Merhaba,

7. ábra: Feladatlap az "ismeretlen nyelv élménye" órához
(Készült a ChatGPT-4 segítségével, 2024. február 14.)

A MID tanárképzés során az MI által generált óratervek segítségével többféle készségfejlesztő feladatot készíthetünk a tanárjelöltek számára. Például **egy-egy generált óratervet átalakíthatunk a hallgatókkal a tanult nyelvtanítási módszereknek megfelelően, elmélyítve ezzel a módszertani ismereteiket, vagy egy adott nyelvtanulói csoport igényeire, különböző nyelvi szintekre is szabhatják őket** gyakorlásként. A hallgatók az MI segítségével a kritikai gondolkodásukat, kreativitásukat és a problémamegoldó képességüket is fejleszthetik, ha például Van der Berg és du Plessis (2023) ötletét kipróbálják. Ők azt javasolják, hogy **a saját**

magunk által készített óratervet töltjük fel, és szisztematikusan menjünk végig a Chat-GPT-vel beszélgetve az óra egyes szakaszain, kritikusan értékelve azt, hogy megvalósulnak-e az óratervben kitűzött céljaink, jól kapcsolódnak-e egymáshoz a feladataink.

Ezenkívül nagy segítséget jelenthet a tanárképző számára az MI a példaként szolgáló csoport- és tanulói profilok elkészítésében, amelyek szintén többféle készségfejlesztő feladat alapjaként szolgálhatnak.

1. Fiatalok Kezdő Csoportja (14-18 év)

- **Korosztály:** 14-18 év
- **Nyelvi Szint:** Kezdő
- **Nemzetiségek:** Német, spanyol, francia, kínai, amerikai
- **Cél:** Alapvető kommunikációs készségek megszerzése, alapvető szókinccs és nyelvtani struktúrák elsajátítása, valamint a magyarországi középiskolai vagy egyetemi tanulmányokra való felkészülés.
- **Csoportlétszám:** 12 fő
- **Tanítási Módszerek:** Interaktív játékok, egyszerű beszélgetések, alapszintű szövegértési és írásgyakorlatok.

8. ábra: Csoportprofil (Készült a ChatGPT-4 segítségével, 2023. szeptember 3.)

3.3. Tananyagkészítés, az órákra való felkészülés

Bár ma már egyre több tankönyv és tananyag érhető el, a viszonylag kis választék és a nagyon változatos tanulói igények miatt a magyar mint idegen nyelv tanárai még mindig sokkal több időt töltenek saját tananyagok, kiegészítő anyagok készítésével, mint más nyelvek tanárai. Így egy óraterv elkészítése is sokszor több időbe telik, mint egy másik nyelv esetében. Ezért a feladatszekvenciák készítése a MID-tanárképzés részét is alkotja, amit szintén megkönnyíthetnek az MI-technológián alapuló alkalmazások. Például **generálhatunk képeket** – a teljesség igénye nélkül – a Midjourney⁶, a DALL-E⁷, a Bing vagy a magyar fejlesztésű Tengr.ai⁸ segítségével. Természetesen ezeknek a képeknek a felhasználásakor is meg kell jelölnünk a forrást, de ilyen módon nagyban lecsökkenhet az eddigi – az órák előtt a megfelelő és jogtiszta képek keresésével eltöltött – idő. A képkészítést mint feladatot is bevihetjük a MID-óra, hiszen izgalmas tevékenység lehet a magyar nyelv gyakorlására a promptok írása. Ebben az esetben azonban lényeges megjegyezni, hogy jelenleg a Bing, az előfizetéses DALL-E és a Tengr.ai képesek értelmezni a magyar nyelvű utasításokat. A Tengr.ai esetében viszont óvatosságnak kell lennünk, mert nincsenek benne olyan erős tartalmi szűrők az illetlen tartalmak ellen, mint a többi képgenerálóban, ezért fel kell hívni a hallgatók figyelmét arra is, hogy ezt gyerekekkel egyelőre semmiképpen se használják. Egyúttal arra is kitérhetünk, mennyire fontos az órán használni tervezett MI-alkalmazások tulajdonságainak megismerése.

⁶ <https://www.midjourney.com/>

⁷ <https://openai.com/index/dall-e-3/>

⁸ <https://tengr.ai/en>

Kezdő nyelvtanárként régen számomra mindig nagy kihívást jelentett a jogtiszta zenék és videók keresése is, amelyeket különböző feladatok alapjaként, inspirációjaként lehet használni a nyelvórákon (pl. forgatókönyvírás zene alapján, predikció, időbeli sorrendiség megvitatása). A **zenék generálása** ma már megoldódott a Suno⁹ segítségével, amely viszonylag jól kezeli a magyar nyelvet is.¹⁰ A realiztikus **videók generálása** pedig hamarosan például az OpenAI Sora¹¹ nevű programjának köszönhetően válhat valóra. A **hanghatások készítése** is jelentősen egyszerűbbé vált a generatív MI-eszközöknek köszönhetően. Bár jelenleg csak angol nyelvű utasítások megadásával, de könnyen használható például a MyEdit¹² vagy az Artist.io¹³ ezeknek a feladatoknak az elvégzésére. Az így készült hanghatások sokféle beszéd- és íráskészség-fejlesztő vagy projektfeladat alapjaként is szolgálhatnak.

Egy másik probléma, amellyel a magyar mint idegen nyelv tanárai időnként szembesülnek, a nyelvkönyvekhez tartozó hanganyagok hiánya vagy olyan adattárolón (pl. kazetta, CD) történő tárolása, amely ma már nehezen olvasható. Továbbá az is felmerülhet problémaként, hogy ezek digitalizálása sokszor nem megoldható a túlzott használatból eredő sérülések miatt. A legtermészetesebb megoldás, ha magunk olvassuk fel és rögzítjük újra a szöveget, vagy megkérünk erre valakit. Ha erre nincs lehetőségünk vagy időnk, a generatív MI-eszközök is segítségünkre lehetnek. Bár néhány **hanggeneráló MI-alkalmazás** magyarul még kissé mesterségesen, gépiesen hangzik, a hanganyag leiratát felhasználva például a Speechify¹⁴ vagy a Talkpal¹⁵ segítségével felolvastathatjuk a szövegeket. Párbeszédetek esetében nehezebb a dolgunk, mert többféle hang generálására nem minden program alkalmas jelenleg, de például a Kukarellával¹⁶ felolvastathatunk dialógusokat¹⁷ is. Végül pedig a hanggeneráló alkalmazásokkal kapcsolatban érdemes lesz figyelni a ChatGPT fejlődését, mert legújabb változata már beszél, és érzelmi intonációval is rendelkezik, amit már magyarul is próbál használni.¹⁸

Egy-egy ilyen program működésének megismerése nagyban lerövidítheti az órákra való felkészülés idejét. A kép-, zene-, hang- és videógenerátorok segítségével megspórolt időt pedig a hallgatók, kezdő tanárok tanári készségeik fejlesztésére vagy óráik további színesítésére fordíthatják.

A **generatív MI-alkalmazások** viszonylag jó minőségű **magyar szövegeket képesek előállítani** az óráinkhoz, azonban ezek a szövegek nem minden esetben hangzanak természetesnek, a minőségük szoros összefüggésben van a GPT verziójával. Ezért mindig meg kell néznünk, hogy éppen melyikkel dolgozunk. Az újabb modellek sokkal természetesebben hangzanak magyarul is, az ingyenes változat használatát követően azonban a legtöbb esetben még szerkesztésre van szükség. A szöveggenerálók jól alkalmazhatóak például különböző sablonok (motivációs levél, kérvény stb.) vagy egy szöveg variációinak elkészítésére is. Itt érdemes megemlíteni a magyar nyelvre optimalizált nagy nyelvi modelleket, amelyek közül a legújabb fejlesztés a PULI: a BERT-Large, GPT-2 és GPT-3 nyelvmodelleken alapul, és demo GPT-3 változata már mindenki számára elérhető.¹⁹ (Yang et al. 2023: 247–262). A további fejlesztések

⁹ <https://suno.com/>

¹⁰ A korábbi példához kapcsolódva ezen a linken meghallgathatjuk az Anyám tyúkja c. vers (vagy népdal) egy megzenésített változatát: <https://suno.com/song/a2c0da92-bb68-4f06-8adb-b1e3d3c8cc3d>

¹¹ További információ: <https://openai.com/index/sora/>

¹² <https://myedit.online/en/audio-editor>

¹³ <https://artist.io/>

¹⁴ <https://speechify.com/ai-audio/>

¹⁵ <https://app.talkpal.ai/chat>

¹⁶ <https://www.kukarella.com/>

¹⁷ Egy ChatGPT-4o által angolul generált, magyarra fordított bolti dialógust a Kukarella két magyar hangja olvas fel ebben a videóban: https://youtu.be/83Ced88BrNA?si=RAx1KSej7_QT8YVc

¹⁸ Az előző ChatGPT-4o által generált dialógust a ChatGPT-4o olvassa fel ebben a videóban: https://youtu.be/oSpUjtX0pM?si=NNUI9Nj98gN2S_SF

¹⁹ <https://juniper.nytud.hu/demo/puli>

után várhatóan ezeket a modelleket hatékonyabban fogjuk tudni alkalmazni a magyar nyelvű szövegek generálásához.

Tanárjelöltként, kezdő nyelvtanárként, anyanyelvi beszélőként kihívást jelenthet a szövegek nyelvi szintjének meghatározása, illetve a nyelvi szintekhez való igazítása. Ezért nagy a kísértés, hogy a szöveggenerálókat használjuk erre. Ismét meg kell jegyeznünk, hogy az MI-alkalmazások sajnos nem alkalmasak erre a feladatra. A szövegek színtezéséhez a legnagyobb segítséget ma a MagyarOK nyílt korpuszának²⁰ használata jelentheti. Az, hogy ez pontosan hogyan történhet, azonban már egy másik tanulmány tárgyát képezhetné.

Végül pedig térjünk rá **a már meglévő szövegekhez MI-alkalmazások segítségével történő feladatok készítésére!** Először a szövegek esetében a szerzői jogokat kell megvizsgálnunk, és ahogy már korábban említettük, védett vagy személyes adatokat tartalmazó szövegeket nem ajánlott feltölteni, mert nem tudjuk pontosan, mi történik velük ezután. Abban az esetben, ha írtunk egy saját szöveget, vagy találtunk egy szabadon felhasználhatót, az MI segítségével különböző típusú feladatokat készíthetünk hozzá, például igaz–hamis állításokat, többválaszos kvízeket, kiegészítendő kérdéseket is írathatunk velük, s egy-egy szöveget akár hiányos szöveggé is alakíthatunk. Ugyanakkor ezekre a generált feladatokra is kritikával kell tekintenünk, minden esetben ellenőriznünk kell a helyességüket. Meg kell ismernünk az adott MI-alkalmazás korlátait. A ChatGPT-4 például megbízhatóbb, amikor kiegészítendő kérdéseket ír egy szöveghez, mint amikor igaz–hamis feladatokat készít hozzá (Berejnoi 2024).

Minden generatív MI-eszköz alkalmazásakor rendkívül lényeges a megfelelő felhasználói utasítás, vagyis prompt megfogalmazása. Az úgynevezett *prompt engineering*-et, vagyis a hatékony utasítások elkészítésének tudományát ma már tanfolyamokon is oktatják, sőt találhatunk olyan weboldalakat is, ahol előre megfogalmazott utasításokat vásárolhatunk. A jól megfogalmazott utasítás ugyanis nagyban megnövelheti az MI-eszközök hatékonyságát. Ezért fontos lenne **a promptírás gyakorlásának bevezetése a MID-tanárképzésbe** is. Az, hogy melyik MI-eszköz milyen formátumú és tartalmú utasításra reagál a legjobban, magától a modelltől függ, ugyanakkor van néhány irányvonal, amelyet érdemes követnünk. A lenti táblázatban összefoglaltam ezeket (1. táblázat). Mivel a legtöbb generatív MI-alkalmazás angol nyelvre optimalizált, a promptírás is ezen a nyelven működik a leghatékonyabban. Bizonyos esetekben ezért érdemes az utasítást angol nyelven megírni – belefoglalva, hogy a válaszadás magyarul történjék.

	Komponens	Szerep	Leírás
1.	Feladat	A cél meghatározása	A jól meghatározott, megfelelően fókuszált feladatleírás a leglényegesebb része az utasításnak.
2.	Szerep	Az interakciók alakítása	A megfelelő koherencia és relevancia érdekében határozzuk meg a modell szerepét (pl. szakértő, nyelvtanár).
3.	Kontextus	A háttérinformációk megadása	Az utasításokban a kontextus megadása is lényeges a releváns válasz érdekében.
4.	Iránymutatás	A paraméterek megadása	A paraméterek megadásával még pontosabb választ kaphatunk. Tisztán ismertessük a szabályokat, ugyanakkor ne írjunk az utasításunkba tagadó mondatokat.
5.	A kimenet formátuma	A válaszok strukturálása	Fogalmazzuk bele az utasításunkba a kívánt kimeneti formátumot (pl. lista, táblázat).

1. táblázat: Irányvonalak a hatékony felhasználói utasítások elkészítéséhez. Készült Singh (2023) és Silver–Szőke–Saumell (2024) alapján

²⁰ https://app.sketchengine.eu/#dashboard?corpname=preloaded%2Fmagyarok_hp2

A kész szövegekkel való munkával kapcsolatban még egy pontot lényeges megjegyezni. Az MI-eszközök képesek arra, hogy a feltöltött szövegekből kinyerjék a gyakori szavakat, kifejezéseket és kulcsszavakat. Például a ChatGPT-4o a *Sketch Engine* korpuszelemző programhoz hasonlóan TF-IDF (*Term Frequency-Inverse Document Frequency*) algoritmust alkalmaz erre a célra (OpenAI modellek 2024). Ugyanakkor jelenleg a korpuszelemző eszközök megbízhatósága ezen a területen lényegesen magasabb. Ezért a gyakori szavak, kulcsszavak és kifejezések azonosításához célszerűbb a *Sketch Engine* használata, például a MagyarOK nyílt korpusának referenciakorpusként való alkalmazásával. Ez a téma szintén önmagában is egy külön tanulmány tárgyát képezhetné.

3.4. Mérés és értékelés

Utolsó gondolatként néhány mondat erejéig kitérek a nyelvórákon alkalmazott mérési és értékelési módszerek tanításának szerepére a MID tanárképzés keretében. Az MI-eszközök elterjedésének kezdete óta számos nyelvtanártól hallottam, hogy nehéz ma már a tanulók íráskészségét értékelni, nem lehet fogalmazást adni házi feladatként, mert megíratják a ChatGPT-vel. Sokan arra figyeltek egy ideig, hogyan lehet detektálni a szöveggenerálók által létrehozott esszéket. Úgy gondolom, az, hogy a tanulók a házi feladat megírására használják az MI-t, természetes dolog, hiszen mindnyájan próbáljuk kihasználni ennek az új technológiának az előnyeit, és ezt a nagyfokú tanulói érdeklődést akár a saját tanári céljaink elérésére is használhatjuk. Ugyanakkor ez rávilágít arra is, hogy milyen nagy szükség van arra, hogy a tanárképzésben lévő hallgatók több más, alternatív, akár **MI által támogatott mérési és értékelési módszerrel** (vö. Khan 2024: 177–197) is megismerkedjenek, amelyek a mérés és értékelés mellett akár a tanulók autonómiájának, felelősségvállalásának és nyelvi készségeinek fejlesztéséhez is hozzájárulhatnak. Fontos azt is megjegyezni, hogy habár nagy a kísértés arra, hogy a tanulók ChatGPT-vel írt szövegeit válaszul mi is a ChatGPT-vel értékeljük ki, ez sajnos nem a legmegfelelőbb megoldás, ugyanis az MI-eszközök ma még nem alkalmasak a helyesírási és nyelvhelyességi hibák megtalálására.

4. Lezárás

A tanulmány megírásának célja egy modern és aktuális kérdéskör feldolgozása volt, amely szorosan kapcsolódik egy olyan témához, amely Nádor Orsolya számára mindig is kiemelkedő jelentőségű volt. A mesterséges intelligencia rendkívül izgalmas és szerteágazó terület, amelyről számos megközelítésből beszélhetünk, így nagy kihívást jelent mindezt egy meghatározott hosszúságú írásban összefoglalni. A jelen tanulmányban igyekeztem minden olyan aspektust érinteni, amely a magyar mint idegen nyelvi tanárképzés, illetve a kezdő MID-tanárok szempontjából hasznos lehet. Összegzésképpen álljon itt néhány tanács az MI eszközök használatával kapcsolatban:

- ❖ Mindig ellenőrizzük, hogy melyik verzióval dolgozunk! Az ingyenes változatok általában nem annyira megbízhatóak, mint a legújabb, előfizetést igénylők.
- ❖ Az MI-alkalmazások által tényként közölt információkat mindig ellenőrizzük!
- ❖ Jól fogalmazzuk meg a felhasználói utasításokat!
- ❖ A felhasználói utasításokban ne adjunk meg személyes adatokat, ne töltsünk fel szerzői jog által védett dokumentumokat!
- ❖ Ne adjunk meg nyelvi szintet! Az alkalmazások még nem képesek a KER gyakorlati alkalmazására. Nem képesek a nyelvi szintet megfelelően meghatározni vagy egy szöveget egy adott nyelvi szinthez igazítani.
- ❖ Ne használjuk helyesírás és nyelvhelyesség ellenőrzésére! Még nem alkalmasak erre (magyarul).
- ❖ Egy MI-alkalmazás órán történő használata előtt mindig ismerjük meg az adott program előnyeit és korlátait!
- ❖ Az MI által készített feladatok megoldásait mindig ellenőrizzük!

Irodalom

- Berejnoi, Daniel 2024. *AI-Powered ESL: Revolutionizing Language Learning for the Digital Age*. Előadás az Innovate ELT konferencián. Barcelona, 2024. május 26.
- Bowen, José Antonio – Watson, C. Edward 2024. *Teaching with AI: a practical guide to a new era of human learning*. Johns Hopkins University Press.
- Chan, Cecilia Ka Yuk – Colloton, Tom 2024. *Generative AI in higher education: the ChatGPT effect*. Routledge.
- Horváth László 2023. Feltáró szakirodalmi áttekintés a mesterséges intelligencia oktatási használatáról. *Pannon Digitális Pedagógia* (ISSN 2786-2445) 3. évf. (2023). 1. sz. DOI: doi.org/10.56665/PADIPE.2023.1.1, 5–17. file:///C:/Users/User/Downloads/5_PADIPE+2023_1_1_Horv%C3%A1th.pdf (2024.06.10.)
- Hyatt, David F. – Beigy, Anne 1999. Making the Most of the Unknown Language Experience: Pathways for reflective teacher development. In: *Journal of Education for Teaching: International research and pedagogy*, 25:1, 31–40.
- ISTE (é.n.). *Bringing AI to School: Tips for School Leaders*. https://cms-live-media.iste.org/Bringing_AI_to_School-2023_07.pdf (2024.06.10.)
- Khan, Salman 2024. *Brave new words: how AI will revolutionize education (and why that's a good thing)*. Penguin Random House LLC. Viking.
- Kostadinovska-Stojchevska, Bisera – Popovikj, Ivana 2019. Teacher Talking Time Vs. Student Talking Time: Moving from Teacher-centered Classroom to Learner-centered Classroom. In: *The International Journal of Applied Language Studies and Culture* (Vol. 2, Issue 2, pp. 25–31). London College of United Knowledge. <https://doi.org/10.34301/alsc.v2i2.22>
- Luckin, Rose – George, Karine – Cukurova, Mutlu 2022. *AI for School Teachers*. CRC Press, Taylor & Francis Group.
- Nádor Orsolya 2017. A magyar mint idegen nyelv tanárok a tanárképzés hullámvasútján. *THL2 2017*: 1–2: 46–56.
- Nádor Orsolya 2018. *Tanított (anya)nyelvünk*. Károli Gáspár Református Egyetem – L'Harmattan Kiadó, Budapest.
- NVIDIA é.n. *What is Generative AI?* <https://www.nvidia.com/en-us/glossary/generative-ai/> (2024.06.08.)
- OpenAI 2024. ChatGPT (4.0 és 4o verzió) [Szoftver]. <https://openai.com/>
- OpenAI modellek 2024. Web: <https://platform.openai.com/docs/models/gpt-4o> (2024.06.10.)
- Porter, Alexis 2023. *Unveiling 6 Types of Generative AI*. <https://bigid.com/blog/unveiling-6-types-of-generative-ai/> (2024.06.10.)
- Shah, Priten 2023. *AI and the future of education: teaching in the age of artificial intelligence*. John Wiley & Sons.
- Silver, Jasmin – Szőke Johanna – Saumell, Vicky 2024. *Generative AI Idea Pack for English language teachers*. Cambridge University Press & Assessment, Cambridge. <https://bit.ly/4cbllhk> (2024.06.10.)
- Singh, Vikash 2023. *Crafting Effective AI Prompts: A Guide to Elevate your Conversation*. <https://medium.com/@vikashsinghy2k/crafting-effective-ai-prompts-a-guide-to-elevate-your-conversations-6bbcd26b9b3c> (2024.06.10.)
- van den Berg, Geesje – du Plessis, Elize 2023. ChatGPT and Generative AI: Possibilities for Its Contribution to Lesson Planning, Critical Thinking and Openness. In: *Teacher Education. Education Sciences*, 13(10), 998. doi:10.3390/educsci13100998.
- Yang Zijian Győző – Dodé Réka – Ferenczi Gergő – Héja Enikő – Jelencsik-Mátyus Kinga – Körös Ádám – Laki László János – Ligeti-Nagy Noémi – Vadász Noémi – Váradi Tamás 2023. Jönnek a nagyok! BERT-Large, GPT-2 és GPT-3 nyelvmODELLEK magyar nyelvre. In: *XIX. Hungarian Computational Linguistics Conference*, 2023, Szeged, Hungary. https://rgai.inf.u-szeged.hu/sites/rgai.inf.u-szeged.hu/files/mszny2023_0.pdf (2024.06.10.)

Vermeki, Boglárka

The possibilities and (current) limitations of using generative artificial intelligence in Hungarian as a foreign language teacher training

The rapid development of artificial intelligence (AI) has significantly transformed our everyday lives, making it unsurprising that it has also revolutionized language education. This paper delves into the potential and current limitations of generative AI applications within the Hungarian as a Foreign Language teacher training program. The objective is to offer a comprehensive overview of the primary features of generative AI and the various aspects that need to be considered when implementing these tools in teacher training programs and educational settings. It emphasizes how the integration of AI tools can enhance teachers' instructional strategies and modernize lesson planning and the preparation process. Additionally, this paper also highlights the importance of fostering critical analytical skills among teachers to evaluate and utilize AI tools effectively.