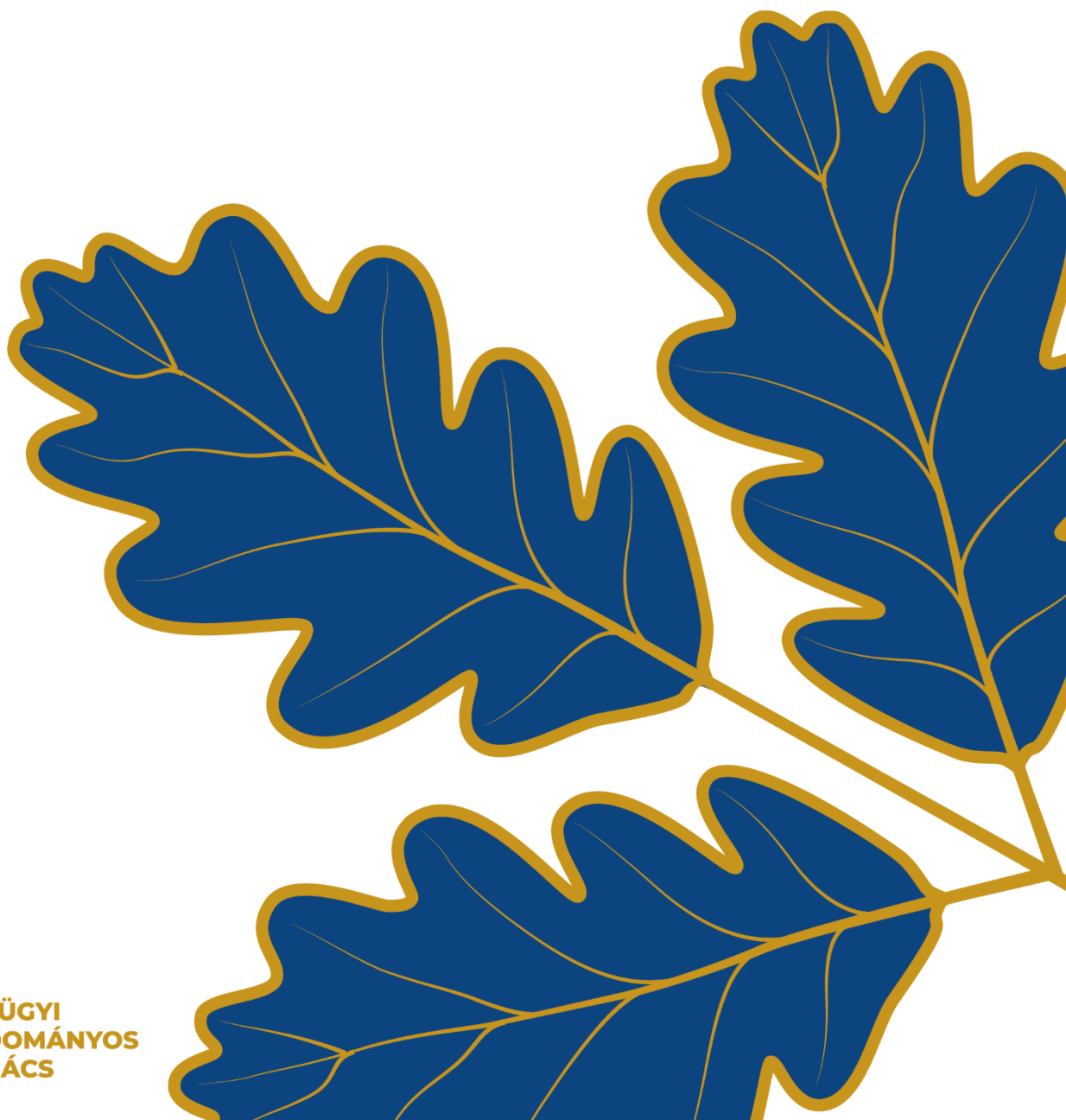




XIII. évfolyam 2024/I. szám



RENDVÉDELEM





B U D A P E S T

– 2024 –

A BELÜGYI TUDOMÁNYOS TANÁCS

ONLINE FOLYÓIRATA

IMPRESSZUM

KIADJA:

Belügyminisztérium

FELELŐS KIADÓ:

Dr. Felkai László közigazgatási államtitkár, Belügyi Tudományos Tanács, elnök

FELELŐS SZERKESZTŐ:

Dr. Sabjanics István ny. r. dandártábornok, Belügyi Tudományos Tanács, ügyvezető alelnök

KIADÓ:

Cím: 2090 Remeteszőlős, Nagykovácsi út 3.

Postacím: 1903 Budapest, Pf. 314.

E-mail: titkar@bm-tt.hu

Telefon: +36 26 795 911, BM: 24-592

Fax: +36 26 795 919, BM: 24-650

HU ISSN 2560-2349

Tartalom

Szerkesztői előszó	I
Papp Orsolya: A „Rendvédelmi jog és közigazgatás I.” című okostankönyv módszertani bemutatása.....	3
Végh János: Mi is valójában a mesterséges intelligencia?	25
Bezerédi Imre: Komplex rendvédelmi MI politikai stratégiai és technológiai javaslatok a ChatGPT és más LLM-ek szempontjából.....	41
Erdélyi Dóra- Leiter Miklós Patrik: Kérdezz bármit, tudom a választ! – avagy a ChatGPT adatvédelmi kérdései, különös tekintettel a személyes adatok védelmére és a kockázati besorolásra	64
Szilágyi Róbert: A ChatGPT és más nagy nyelvi modellek (LLM-ek) alkalmazásának lehetőségei kis- és közepes vállalkozásokban.....	76
Péterfi Orsolya: Nagy nyelvi modellek gyógyszeripari alkalmazhatóságának lehetőségei.....	86
Szerzők	104

Szerkesztői előszó

A Rendvédelem folyóirat XIII. évfolyam 2024./1-2. számában dr. Felkai László, a Belügyminisztérium közigazgatási államtitkára, a Belügyi Tudományos Tanács elnöke által 2023 áprilisában kiírt *A ChatGPT és más nagy nyelvi modellek alkalmazásának lehetőségei és egyben biztonsági kérdései* című pályázati felhívásra érkezett egyes pályaművek átdolgozott tanulmányai olvashatóak.

Napjainkban, a tudásalapú, információs társadalom korában a társadalom tagjai egy újfajta ipari forradalom részesei, amelynek ténye sok esetben nem is tudatosul a hétköznapi emberekben. Új jelenségként a világ köztudatában megjelent ChatGPT egyenesen „felrobbantotta” a világ információs társadalmának jelentős részét. Olyan radikálisan új jelenségről beszélünk, amely képes volt az indulását követő ötödik napon már több mint 1 millió felhasználót felmutatni, majd néhány hónap alatt több mint 100 milliós nagyságrendet elérni, és a regisztráltak száma folyamatosan növekszik. A ChatGPT és más nagy nyelvi modellek (*Large Language Model, LLM*) használata széles társadalmi rétegeket érint, az élet minden területére beszivárog, kihatással van az egyének magánéletére, a társadalmi szervezetek, a gazdasági és állami szereplők tevékenységére, továbbá a társadalmi entitások közötti kapcsolatokat, folyamatokat is mélyrehatóan átalakíthatja, jelentős és számos etikai és jogi kérdést felvetve.

Az információs társadalom jelenlegi fejlődési szakaszában közép- és hosszútávon azok a társadalmak lesznek képesek hatékonyan és eredményesen működni, valamint egymással együttműködni, melyek a természeti erőforrásokat költséghatékonyan hasznosítják, a társadalmi és gazdasági folyamatokat pontosan elemzik, értékelik, az új tudományos eredményeket innovatívan, pragmatikusan hasznosítják.

Az információtömeg robbanásszerű növekedése a Belügyminisztérium irányítása alá tartozó rendészeti-, egészségügyi-, köznevelési-, gondoskodáspolitikai-, vízügyi- és önkormányzati szektorra is jellemző. Ugyanakkor az információk mennyisége, összetettsége, valamint a rendelkezésre álló humán erőforrás miatt az adatok hatékonyan már nem feldolgozhatók, így a bennük rejlő potenciál teljes spektruma nem kiaknázható. E jelenség hátrányos következményeinek csökkentése érdekében elengedhetetlenül szükséges korunk egyik legnagyobb vívmányának és legintenzívebben fejlődő területének, a mesterséges intelligenciának fejlesztésében és alkalmazásában rejlő

lehetőségeket tanulmányozni és előnyeit, hátrányait vizsgálni. A rendészet területén már alkalmaznak működő – jellemzően hibrid, tehát különböző technológiákat ötvöző – kísérleti MI megoldásokat, azonban ezek konkrét rendvédelmi feladatokhoz és jogszabályi keretekhez történő adaptálása kihívást jelent.

Prognosztizálható, hogy az MI megoldások, mint a ChatGPT és más nagy nyelvi modellek, az információs társadalmak életében már a közeli jövőben kiemelt szerepet fognak betölteni, így védelmükre is kiemelt figyelmet kell fordítani. Egyrészt az általuk betöltött funkciók alapján lehetnek elsődleges célpontok, illetve az interdependencia jelensége okán a kieső funkció hatásából eredően másodlagos célpontok. Az MI megoldásokra (infrastruktúrák és rendszerek) más országok kiberhadviselési rendszerei, a kiberbűnözés és a kiberterrorizmus tekinthet célpontként, de meg kell említeni az ipari kémkedést is, amely megjelenhet az új, vagy újszerű kutatási eredmények megszerzése kapcsán. Az előzőekre tekintettel a védelemre az MI rendszerek kialakításának és üzemeltetésének minden szakaszában figyelmet kell fordítani.

A potenciális MI kutatási irányok mindegyikére igaz, hogy az MI megoldások gyakorlati megvalósításán túl a jogszabályi keretek kidolgozása is alapos munkát igényel. Ez kiemelten igaz a társadalmi folyamatokba, a társadalmi környezetbe aktívan beavatkozó megoldások (például az önvezető járművek, a ChatGPT és más nagy nyelvi modellek) szabályozására, mert itt nem meglévő jogszabályokat kell módosítani, hanem olyan, eddig nem kidolgozott elveket kell megalapozni, megalkotni és lefektetni, melyek a társadalmat kiemelten foglalkoztathatják. Ennek okán szükséges megvizsgálni, hogy az MI kapcsán értelmezhető autonómiának, illetve adaptivitásnak mekkora szabadságfokot lehet engedni, és milyen prioritásokat kell megfogalmazni az adott jogszabályi keretek között.

A ChatGPT és más nagy nyelvi modellek alkalmazásának lehetőségei és egyben biztonsági kérdései című pályázat meghirdetésének célja az volt, hogy a Belügyi Tudományos Tanács az állomány tagjainak szakmai ismereteit és értékes gyakorlati tapasztalatait, valamint a téma iránt érdeklődő szakemberek véleményét és javaslatait megismerje. Olyan pályaműveket vártunk, amelyek elsősorban az életszerű, valós problémákra megoldást kínáló ötletek rendszerszemléletű megközelítéssel történő megoldásait mutatják be, és

nem műszaki és matematikai megközelítésű tartalmat tükröznek.

A pályázati felhívásra a feltöltési határidőig, 2023. október 1. 24:00 óráig összesen 26 érvényes jeligés pályamű érkezett a Belügyi Tudományos Tanács online pályázati felületére.

A bírálati folyamat (az elbírálási határidő 2023. november 15. volt) két részből állt. Az első részben az online pályázati felületen a téma tudományos és szakmai elismertségű szakértői bírálóként pontozták és szövegesen értékelték a benyújtott jeligés, anonim pályázatokat. Ezt követően a felkért bizottság a bírálók értékelése alapján döntött a díjazásról.

Jelen XIII. évfolyam 2024./1. számot 2024 tavaszán a 2024/2. szám követi, melyben szintén a pályázati felhívásra benyújtott pályaművek közül a szerkesztőség döntése alapján kiválasztott, átdolgozott tanulmányok jelennek meg. A publikációk szakmai tartalmáért, naprakész információiért felelősséget nem vállalunk.

E szám első cikke nem tartozik a pályaművek közé, de a digitális oktatásra való átállás aktualitása miatt fontosnak tartjuk megjelentetését: Papp Orsolya a *Rendvédelmi jog és közigazgatás I.* című szakképzési tananyag (jegyzet) digitális tartalmakkal történt bővítését és annak pedagógiai módszertani alkalmazását, okostankönyvét mutatja be. Ezt követően négy, *A ChatGPT és más nagy nyelvi modellek alkalmazásának lehetőségei és egyben biztonsági kérdései* című pályázatra benyújtott, átdolgozott pályaművet közlünk: Végh János, Bezerédi Imre, Erdélyi Dóra-Leiter Miklós Patrik, Szilágyi Róbert és Péterfi Orsolya publikációját.

Az értékes és változatos témájú publikációkhoz hasznos olvasást kívánunk!

Dr. Sabjanics István ny. r. dandártábornok
felelős szerkesztő

PAPP ORSOLYA

A „Rendvédelmi jog és közigazgatás I.” című okostankönyv módszertani bemutatása

The „Law Enforcement Law and Public Administration I.” methodological presentation of the smart textbook

DOI: [HTTPS:// DOI.ORG/10.53793/RV.2024.I.I](https://doi.org/10.53793/RV.2024.I.I)

Absztrakt

A Z generáció oktatása a 21. századi pedagógusok egyik legnagyobb kihívása. Számukra természetes a digitális technológiák használata, figyelmük rövid ideig tartható fenn, a szükségesnél több energiát nem fektetnek tevékenységeikbe, minden kérdésre azonnal választ szeretnének kapni. A cikk a KÖFOP-2.1.5–VEKOP-16-2016-00001 azonosítószámú, „A versenyképes közszolgálat személyzeti utánpótlásának stratégiai támogatása” című projekt keretében készült Rendvédelmi jog és közigazgatás I. című szakképzési tananyag (jegyzet) digitális tartalmakkal történt bővítését és annak pedagógiai módszertani alkalmazását mutatja be. A kiegészítő digitális tartalom támogatja a szemléltetést, segít az összefüggések felismerésében, épít a diákok aktivitására a feldolgozási folyamatban, valamint lehetőséget ad az önszabályozó tanulásra.

KULCSSZAVAK: OKOSTANKÖNYV, Z GENERÁCIÓ, DIGITÁLIS TARTALOM, JÁTÉKOS ELEMELK, INTERAKTÍV FELADATOK

Abstract

Educating Generation Z represents one of the foremost challenges for 21st-century educators. Their inclination toward using digital technologies is natural, with attention spans that are short-lived, and a tendency not to invest more energy than necessary in activities; they seek immediate answers to every question. This article introduces the expansion of the vocational training material titled "Rendvédelmi jog és közigazgatás I." (Law Enforcement Law and Public Administration I.), developed within the framework of the project titled "Strategic Support for the Competitive Personnel Supply of Public Service" with identification number KÖFOP-2.1.5–VEKOP-16-2016-00001. The expansion involves the incorporation of digital content and illustrates its pedagogical methodological application. The supplementary digital content supports visualization, aids in recognizing connections, fosters student engagement throughout the processing phase, and provides opportunities for self-regulated learning.

KEYWORDS: SMART TEXTBOOK, GENERATION Z, DIGITAL CONTENT, GAMIFIED ELEMENTS, INTERACTIVE EXERCISES









Bevezetés

A 2023. évben a Belügyminisztérium Közigazgatási Államtitkárának személyügyi határozata alapján célfeladatként kaptam a pályaoorientációs szakképzés (Közszolgálati technikus és a Rendészeti ör szakmák) Képzési és Kimeneti Követelményeiben meghatározott szakmai követelményekhez illeszkedő Rendvédelmi jog és közigazgatás I. című tansegédlet hatályosítását, digitális tartalommal való kiegészítését.

A Z generációs diákok hatékony oktatása hatalmas kihívás. Pörgősek, figyelni rövid ideig képesek, türelmetlenek, mindent azonnal akarnak, ugyanakkor nem szeretnek a szükségesnél több energiát fektetni tevékenységükbe, okos eszközeik pedig szinte „testrészükhöz”-ként funkcionálnak.

A digitális tartalmak sokféle formátumban kínálnak információkat a tanulóknak, amelyek segítik figyelmük fenntartását és növelik a motivációjukat. Elérésük a tananyagban elhelyezett QR kódok segítségével lehetséges. Tehát a QR kód digitális tartalmat rejt, amely pdf formátumban kattintással (jobb egérgomb - Helyi menü - Hivatkozás megnyitása új lapon), nyomtatott formában pedig telefonon a QR kód olvasó alkalmazás használatával nyitható meg.

A QR kód színes kerete a digitális tartalom fajtájára utal, melyről a Tartalomjegyzék után helyet kapott **Digitális kalauz** ad tájékoztatást (i. sz. kép).

A keret színe	A digitális tartalom fajtája
	Előadás videó
	Interaktív előadás videó
	Gyakorló feladat (különböző típusú)
	Játék
	Interaktív kép
	(Interaktív) gondolattérkép
	Infografika (hangos, animált), Sway, Diasorozat, Ppt
	Egyéb tartalom (kapcsolódó videó, weboldal, fájl stb.)

i. sz. kép: Digitális kalauz

Forrás: Képernyőkép, szerzői felvétel

A digitális tartalomnak számos előnye van: sokkal inkább „vonzza” a diákok figyelmét, érdekessé teszi a tananyagot, ezáltal is növelve a tanulók motivációját és tanulási eredményességét. Mobilbarát, internetkapcsolaton keresztül okos eszközzel bárholnan, bármikor elérhető, böngészőben megoldható. Nem szükséges a nyomtatott tankönyv hozzá, illeszkedik a Z generáció életstílusához, könnyen integrálható a mindennapi technológia-használatba.

Az alkalmazások megválasztásánál fontos szempont volt, hogy magyar nyelvű felülettel rendelkezzen, valamint az elkészített tartalom nyilvánosan ne legyen kereshető, csak adott link birtokában legyen elérhető.

A tananyaghoz öt darab, fejezetenként rendezett Wakelet gyűjtemény készült, amelyből az összes digitális tartalom elérhető. A Wakelet gyűjtemények QR kódjának kerete szürke színű.

Az okostankönyv fejezetei és Wakelet gyűjteményei

1. [Jogi alapismeretek](#)

A fejezetben a diákok megismerkednek a magyar jogrendszer felépítésével, a jogforrások fajtáival és létrejöttük folyamatával (azaz egy elgondolásból, hogyan lesz mindenki számára kötelezően alkalmazandó szabályozás), valamint a jogalkalmazás alapjaival (2. sz. kép).

A gyűjtemény 42 elemet tartalmaz.



2. sz. kép: Wakelet gyűjtemény
Forrás: Képernyőkép, szerzői felvétel

2. [Állam-, Alkotmány-, Nemzetközi jogi és Munkavédelmi alapismeretek](#)

Ez az okostankönyv legterjedelmesebb és legtöbb témát tárgyaló fejezete. A diákok megismerkednek az állam felépítésével, az alkotmányjog szabályozási területével, Magyarország Alaptörvényének szerkezetével és főbb fejezeteivel, az állampolgárság fogalmával és szabályozásával, valamint képet kapnak az emberi jog, a nemzetiségi jog, a humanitárius jog, a közösségi jog (Magyarország és az Európai Unió) jogterületeiről. A fejezet vége a munkavédelem kérdéskörét tárgyalja (3. sz. kép).

A gyűjtemény 63 elemet tartalmaz.



3. sz. kép: Wakelet gyűjtemény
Forrás: Képernyőkép, szerzői felvétel

3. [Polgári jog alapjai](#)

A fejezet röviden bemutatja a polgári jogi szabályozás kódexét, a Ptk.-t, annak szerkezeti felépítését és tartalmi tagozódását (4. sz. kép).

A gyűjtemény 27 elemet tartalmaz.



4. sz. kép: Wakelet gyűjtemény
Forrás: Képernyőkép, szerzői felvétel

4. [Büntetőjog – Általános rész](#)

A fejezet a Btk. Általános részében foglaltakat tárgyalja, a diákok megismerkednek a bűncselekmény fogalmával, ismérveivel, tényállási elemeivel és az elkövetői körrel (5. sz. kép).

A gyűjtemény 47 elemet tartalmaz.



5. sz. kép: Wakelet gyűjtemény
Forrás: Képernyőkép, szerzői felvétel

5. [Kriminálisztika](#)

A fejezet témája a kriminálisztika, mint tudományterület, annak tagozódása és felépítése. A diákok megismerkednek a témakör alapfogalmaival is (6. sz. kép).

A gyűjtemény 16 elemet tartalmaz.



6. sz. kép: Wakelet gyűjtemény
Forrás: Képernyőkép, szerzői felvétel

A tananyag digitális elemei folyamatosan frissülnek, követik a jogszabályváltozásokat, illetve folyamatos az interaktív tartalom bővítése is.

Digitális tartalmak

A digitális tartalom élményszerűbbé teszi a tanulást, segíti a könnyebb megértést, az interaktív feladatok a diákokat is cselekvésre ösztönzik. A feladatok kerete sötétzöld színű, a QR kód pedig annál a tananyagrésznél lett elhelyezve, amelynek ismeretanyagára vonatkozik. A feladatok egyaránt alkalmasak tanórai megoldásra és tanórán kívüli felkészülésre, ezzel is támogatva az önszabályozó tanulást.

Az azonos ismeretanyagra építő, de eltérő típusú feladatok differenciálásra is alkalmasak.

Feladatok

A digitális tartalom fontos részét képezik a feladatok, melyeknek nagy része a LearningApps.org oldalon készült. A tankockák a játék-alapú tanítás-tanulás eszközei, az interaktív felület gyakorlási lehetőséget ad a tanulóknak. Segítik az ismeretátadást és az ismeretszerzést, fejlesztik a digitális kompetenciákat, valamint az együtt-tanulás élményét is nyújthatják játékos környezetben. A tankockákba beágyazható szöveges, képi, hang- és videó részletek lehetővé teszik a látványos, élmény alapú oktatást. A sablonok között több játékot is találunk, amelyek segítenek fenntartani az érdeklődést és növelik a diákok motivációját (7. sz. kép).

Az itt készített feladatok megoldása azonnal ellenőrizhető, így a tanuló hamar visszajelzést kap tudásszintjéről, illetve hiányosságairól. Minden feladat korlátlan számban megoldható, ezért ezek segítik az ismeretek elmélyítését is.

Az okostankönyvben összesen 186 darab itt készült elem található, QR kódjának kerete sötétzöld színű.

A LearningApps.org oldal számos lehetőséget kínál a különböző típusú feladatok összeállítására. A következőkben az okostankönyvben szereplő egy-egy feladattípust és a használatával elérhető pedagógiai célokat mutatom be.

Csoportba rendezés

Ennél a tankockánál a megadott szempontok szerint 2 vagy több csoportba lehet rendezni az érkező elemeket (8. sz. kép).

A tankocka sokoldalúan használható feladattípus:

- jellemzők szétválogatása, összehasonlítása (2, 3 vagy 4 szempont szerint)
- speciális esete az igaz/hamis
- nehezebb, összetettebb feladatok készítése.

Elérhető pedagógiai célok:

- A diákok analitikus és kritikus gondolkodásának fejlesztése. A diákoknak meg kell vizsgálnia és elemeznie kell az adott jellemzőt annak érdekében, hogy a helyes csoportba rendezzék.
- A képes feladatoknál meg kell érteniük a jellemzők közötti kapcsolatokat és a logikát, ezáltal vizuális észlelésüket és logikai gondolkodásukat fejleszti.
- A tanult ismeretek begyakorlása.



7. sz. kép: Feladatok a learningapps.org oldalon
Forrás: Képernyőkép, szerzői felvétel



8. sz. kép: Csoportba rendezés
Forrás: Képernyőkép, szerzői felvétel

- A csoportosítás során a diákoknak önállóan kell gondolkodni és döntéseket hozni, ez erősítheti a tanulás iránti motivációt, tét nélkül, bátran kipróbálhatják önmagukat.
- Az igaz/hamis feladatok döntési helyzetbe hozzák a tanulót, célja az ismeretellenőrzés.

A következő feladatban a diákoknak az érkező állításokat kell a megfelelő csoportba húzni, aszerint, hogy az a kriminalisztikára vagy a kriminológiára igaz (9-10. sz. kép).



9. sz. kép: Kriminalisztika vagy kriminológia?
Forrás: Képernyőkép, szerzői felvétel



10. sz. kép: Tankocka elérése
Kriminalisztika vagy
kriminológia?
Forrás: Képernyőkép, szerzői
felvétel

A következő feladatban az érkező elemeket három csoportba kell sorolni (11-12. sz. kép).



11. sz. kép: Büntethetőségi akadályok csoportosítása
Forrás: Képernyőkép, szerzői felvétel



12. sz. kép: Tankocka elérése
Büntethetőségi akadályok
csoportosítása
Forrás: Képernyőkép,
szerzői felvétel

Ezzel a típussal nehezebb feladat is készíthető. Az alábbi feladatban már 4 csoportba kell rendezni az érkező elemeket: az alapvető jogokat tartalmuk szerint kell csoportokba rendezni (13-14. sz. kép).

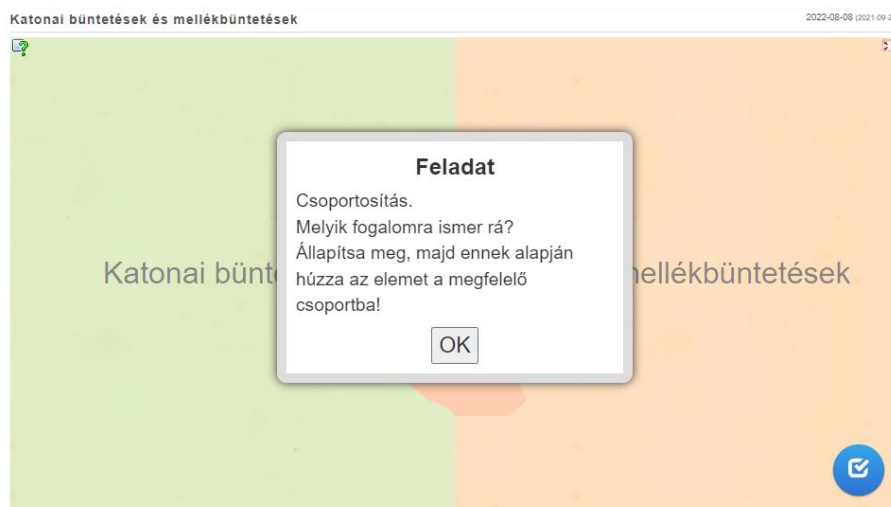


13. sz. kép: Alapvető jogok csoportosítása
Forrás: Képernyőkép, szerzői felvétel



14. sz. kép: Tankocka elérése
Alapvető jogok csoportosítása
Forrás: Képernyőkép, szerzői felvétel

Ebben a formában összetettebb feladat is készíthető, ennek megoldása már bővebb ismeretanyagot és kritikus gondolkodást is elvár a diákoktól. A megjelenő meghatározás alapján rá kell ismerni a megadott fogalomra, majd azt az annak megfelelő csoportba húzni (15-16. sz. kép).



15. sz. kép: Katonai büntetések és mellékbüntetések
Képernyőkép, szerzői felvétel



16. sz. kép: Tankocka elérése
Katonai büntetések és mellékbüntetések
Forrás: Képernyőkép, szerzői felvétel

A nehezebb feladatok közé sorolható, másféle gondolkodást kíván, alaposabb ismereteket vár el a tanulótól az a feladat, amelynél a csoportba rendezés szempontjai között megjelenik a „Mindegyikre igaz” csoport. Azon kívül, hogy el kell dönteni adott jellemző melyik csoportba sorolható, figyelni kell a közös jellemzőkre is. A következő feladat erre mutat példát (17-18. sz. kép).

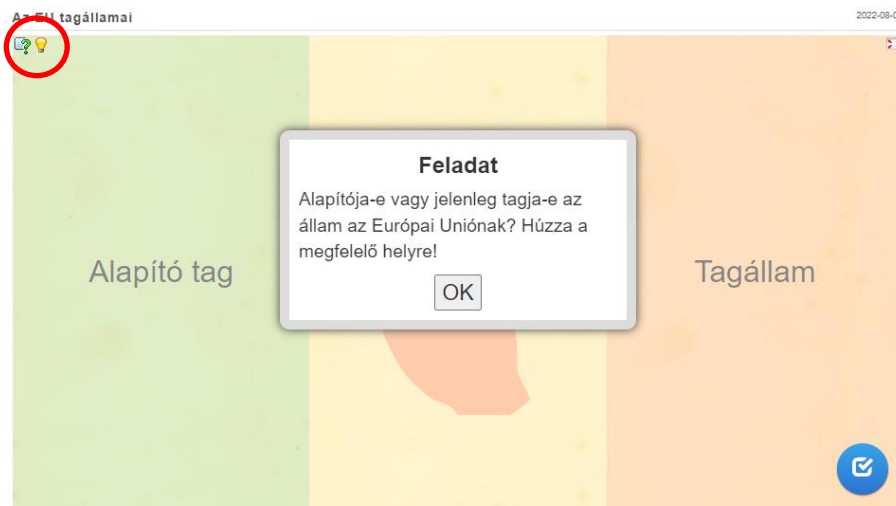


17. sz. kép: A birtokvédelem eszköze
Forrás: Képernyőkép, szerzői felvétel



18. sz. kép: Tankocka elérése
A birtokvédelem eszközei
Forrás: Képernyőkép, szerzői felvétel

Szintén nehezített feladat, amikor a jellemzőket többféle szempontból kell vizsgálni és eszerint a megfelelő csoportba rendezni (19-20 sz. kép).



19. sz. kép: Az EU tagállamai
Forrás: Képernyőkép, szerzői felvétel



20. sz. kép: Tankocka elérése
Az EU tagállamai
Forrás: Képernyőkép, szerzői felvétel

A feladat megoldásához segítség is használható, erre utal a „villanykörte - Súlyó megjelenítése” ikon a bal felső sarokban, amelyre kattintva két weboldalra jut el a diák ([Országismertető](#); [A Európai Unió tagállamai](#)) (19. sz. kép).

Alaposabb ismeretet és annak alkalmazását várja el a következő feladat megoldása a diákoktól: az érkező fogalomról fel kell ismerni, hogy a bűncselekmény melyik elkövetőjét írja le, majd eldönteni, hogy hajt-e végre tényállási elemet, azaz a tettes vagy a részes csoportba sorolható (21-22. sz. kép).



21. sz. kép: Tettes vagy részes?
Forrás: Képernyőkép, szerzői felvétel

A csoportba rendezés speciális esete az igaz/hamis feladattípus, amely döntési helyzetbe állítja a diákot: az érkező állításról el kell dönteni, hogy a megadott szempont szerint igaz vagy hamis, majd a megfelelő csoportba húzni (23-24. sz. kép).



23. sz. kép: Magyarország Alaptörvénye (2011. április 25.)
Forrás: Képernyőkép, szerzői felvétel

A feladatban jellemzők helyet képet is megadhatunk: az adott képen szereplő dokumentumról kell a diáknak eldöntenie, hogy tartalmazhat-e jogszabályt, majd ennek megfelelően a választott csoportba húzni (25-26. sz. kép). A képre kattintva az nagyméretben is megtekinthető.



22. sz. kép: Tankocka
elérése
Tettes vagy részes?
Forrás: Képernyőkép,
szerzői felvétel



24. sz. kép: Tankocka elérése
Magyarország Alaptörvénye
Forrás: Képernyőkép, szerzői
felvétel



25. sz. kép: Jogsabályok
Forrás: Képernyőkép, szerzői felvétel



26. sz. kép: Tankocka elérése
Jogsabályok
Forrás: Képernyőkép, szerzői felvétel

Párkereső

Ez a feladattípus kiválóan alkalmazható fogalmak és definíciók tanulására és rögzítésére.

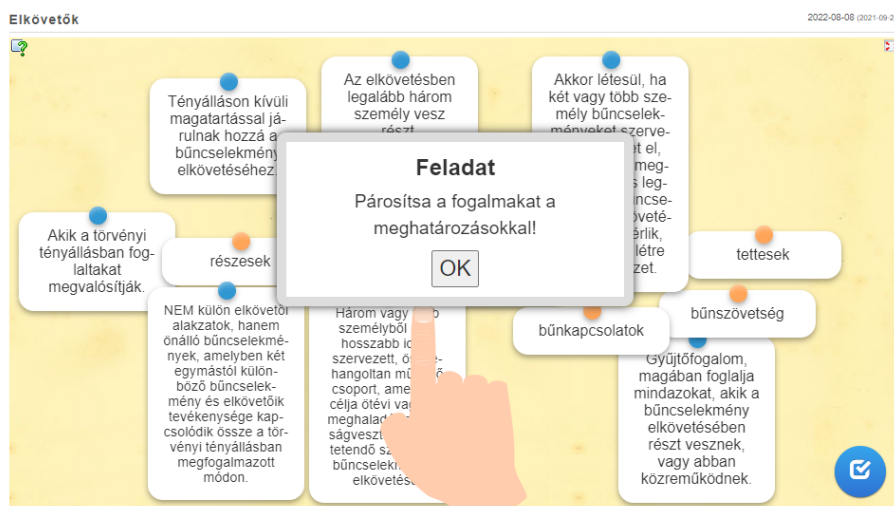
Elérhető pedagógiai célok:

- Elsősegíti az információk összekapcsolását és értelmezését, a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása, szövegértés fejlesztése.

Az előző pontban bemutatam a *Tettes vagy részes?* c. tankonckát, amelynek egyszerűbb változata az *Elkövetők* c. tankocka: a diák feladata az elkövetői körrel kapcsolatos fogalom és meghatározás párosítása. Ez a feladatpáros differenciálásra is alkalmas (28-29. sz. kép).



27. kép Párkereső
Forrás: Képernyőkép, szerzői felvétel.



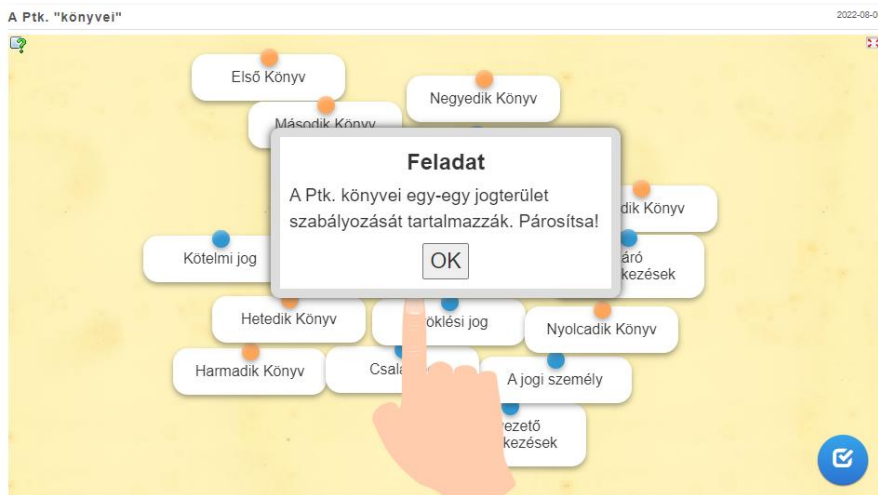
28. sz. kép: Elkövetők
Forrás: Képernyőkép, szerzői felvétel



29. sz. kép: Tankocka elérése
Elkövetők
Forrás: Képernyőkép, szerzői felvétel

Fejlesztési céllal többféle dolgot is lehet párosítani, pl. fogalmat és meghatározást. Erre mutat példát az előző tankocka: kérdés – rövid válasz; kéttagú összetett mondat első és második tagmondata – törvény száma és megnevezése; stb.

A következő tankockában a Ptk. egyes részeinek megnevezését és azok tartalmát kell összerendelni (30-31. sz. kép).



30. sz. kép: A Ptk. „könyvei”

Forrás: Képernyőkép, szerzői felvétel



31. sz. kép: Tankocka elérése
A Ptk. „könyvei”

Forrás: Képernyőkép, szerzői felvétel

Egyszerű sorbarendezés

Ez a feladattípus kiválóan alkalmazható az összefüggések, az elemek közötti logikai kapcsolatok felismerésének gyakoroltatására és a problémamegoldó készség fejlesztésére (32. sz. kép).

Elérhető pedagógiai célok:

- Elsősegíti az információk rendszerezését és értelmezését, a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása.



32. sz. kép: Egyszerű sorbarendezés

Forrás: Képernyőkép, szerzői felvétel

Ezt a feladattípust használtam a jogalkotási folyamat szakaszainak megértetéséhez, ahol a diákoknak az egyes szakaszokat kellett időrendi sorrendbe rakniuk, annak bekövetkezése szerint (33-34. sz. kép).



33. sz. kép: A jogalkotási folyamat szakaszai

Forrás: Képernyőkép, szerzői felvétel



34. sz. kép:
Tankocka elérése
A jogalkotási folyamat
szakaszai

Forrás: Képernyőkép, szerzői felvétel

Idővonal

Ez a feladattípus dinamikus eszköz, kiválóan alkalmazható a folyamatok és más számmal „kifejezhető” információk gyakoroltatására, a problémamegoldó készség fejlesztésére (35. sz. kép).

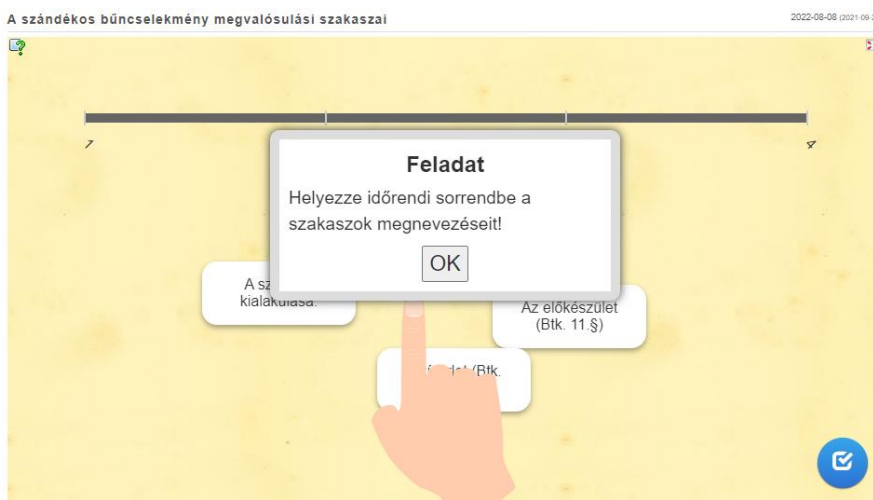
Elérhető pedagógiai célok:

- Szemléletes, vizuálisan is elősegíti az információk rendszerezését és értelmezését, valamint a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása.

Hasonló az előzőekben bemutatott egyszerű sorbarendezés feladathoz, de itt a hangsúly az időbeliségen, azaz a kronológián van. Ez hasznos a folyamatok elsajátításánál is (36-37. sz. kép).



35. sz. kép: Idővonal
Forrás: Képernyőkép, szerzői felvétel



36. sz. kép: A szándékos bűncselekmény megvalósulási szakaszai
Forrás: Képernyőkép, szerzői felvétel



37. sz. kép: Tankocka elérése
A szándékos bűncselekmény megvalósulási szakaszai
Forrás: Képernyőkép, szerzői felvétel

Ez a feladattípus összetett feladat készítésére is alkalmas, amelyben „bármilyen” számmal kifejezhető adat (létszám, szavazatok száma stb.) alapján lehetséges az idővonalon történő elrendezés.

A *Törvényi erő* c. tankockában a diáknak nemcsak bővebb ismeretanyaggal kell rendelkeznie (pl. az országgyűlés minimális létszáma, az egyes jogszabályok elfogadási feltételei stb.), hanem már számolnia is kell, azaz a tanultakat a gyakorlatban is tudni kell alkalmaznia (38-39. sz. kép).



38. sz. kép: Törvényi erő
Forrás: Képernyőkép, szerzői felvétel

A 38. képen látható feladat: „A Covid-19 járvány első hulláma idején az Országgyűlés éppen csak határozatképes, azaz minimális létszámú. A mai ülésen az Alaptörvény egyik módosítását, egy új sarkalatos törvény elfogadását és egy egyszerű törvény hatályon kívül helyezését tárgyalják a napirendi pontok szerint. Helyezze el a nevezett törvényeket a számegyenesen, aszerint, minimálisan hány képviselői szavazat szükséges az elfogadásukhoz! Tipp: Az eger húzásával a pontos értéket be tudja állítani.”

Hiányos szöveg

Ez a feladattípus is nagyon sokoldalú. A feladat megoldása elvárja a diákoktól a szövegben előforduló adatok, megállapítások, érvek, álláspontok azonosítását, értelmezését; a szövegbeli utalások, hivatkozások szerepének, a szöveg és más információk kapcsolatának értelmezését; a szöveg egészének jelentése, jelentésrétegei közötti összefüggések megmutatását; a cím és a szöveg kapcsolatának értelmezését; a szöveg tartalmi-logikai gondolatmenetének, szerkezetének vizsgálatát és a felépítés által közvetített jelentés azonosulását (40. sz. kép).

Elérhető pedagógiai célok:

- Szemléletes, vizuálisan is elősegíti az információk rendszerezését és értelmezését, valamint a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása, elmélyítése.
- Szövegértés fejlesztése, szókinccs bővítése, kritikus olvasási készségek fejlesztése.

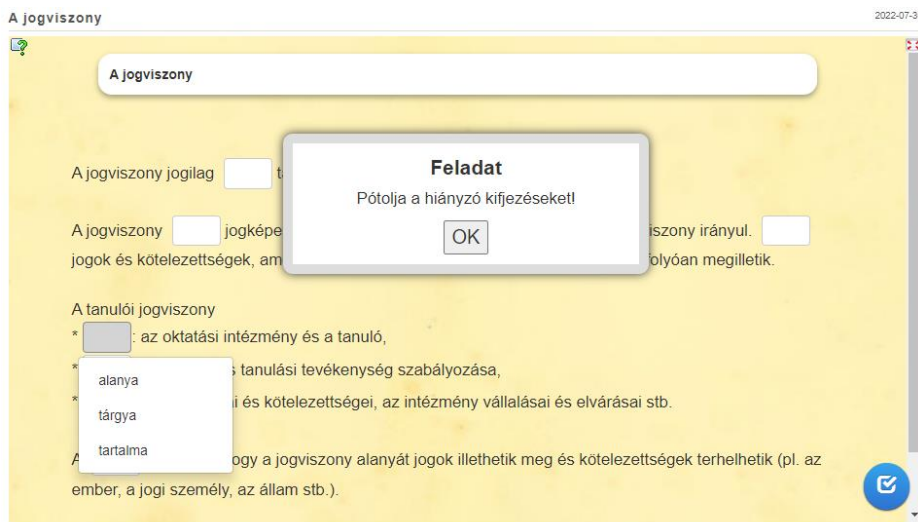
A könnyebb feladatoknál a kiegészítő elem kiválasztása történhet listából, ez segíti a diákot, hiszen ilyenkor a megadott lehetőségek közül kell kiválasztani a helyesnek vélt megoldást (41-42. sz. kép).



39. sz. kép: Tankocka elérése
Törvényi erő
Forrás: Képernyőkép, szerzői felvétel



40. sz. kép: Hiányos szöveg
Forrás: Képernyőkép, szerzői felvétel



41. sz. kép: A jogviszony
Forrás: Képernyőkép, szerzői felvétel

Ez a feladattípus nehezebb feladat készítésére is alkalmas. Ebben az esetben a diáknak saját magától kell beírnia a kiegészítő elemet, ügyelve a helyesírásra és a pontos adatbevitelre. Ezt a variációt alkalmazhatjuk pl. törvényi szöveg vagy fogalmak ismeretének számonkérésekor, ahol segítségként linkelhetjük a törvényi szöveget is.

A következő példában a Btk. XII. fejezet Katonákra vonatkozó rendelkezések elkövetőkre (127.§) vonatkozó szabályozását kell kiegészíteni, segítségként a „villanykörtére” kattintva a hatályos törvényi szöveg olvasható (43-44. sz. kép). Ennek a feladatnak a megoldása a diák szakmai szókincsét is fejleszti.



43. sz. kép: Katonákra vonatkozó rendelkezések
Forrás: Képernyőkép, szerzői felvétel

Ezzel a feladattípussal is támogatható a gyakorlatorientált megközelítés, erre mutat példát a következő feladat (45-46. sz. kép). Ebben a feladatban a törvényi szöveg olvasása közben kell a diákoknak felismerni, hogy a megfogalmazás alapján a diszpozíció melyik fajtájára vonatkozik a példa, majd a helyesnek gondolt megoldást egy listából kell kiválasztaniuk.



42. sz. kép: Tankocka
elérése
A jogviszony
Forrás: Képernyőkép,
szerzői felvétel



44. sz. kép: Tankocka elérése
Katonákra vonatkozó
rendelkezések
Forrás: Képernyőkép,
szerzői felvétel



45. sz. kép: A diszpozíció fajtái (Példák)
Forrás: Képernyőkép, szerzői felvétel

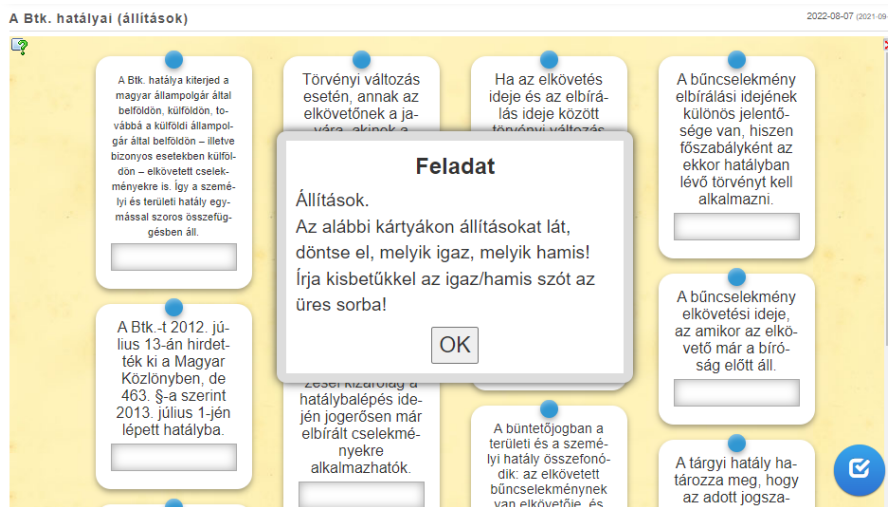
Rövid válasz

Ez a feladattípus is többféleképpen használható: a kis kártyák tartalmazhatnak állításokat, amelyekről el kell dönteni, hogy igazak vagy hamisak; tartalmazhatnak fogalmakat, amiket fel kell ismerni; tartalmazhatnak egyszerűbb gyakorlati feladatokat, amelyek megoldásra várnak (47. sz. kép).

Elérhető pedagógiai célok:

- Szemléletes, vizuálisan is elősegíti az információk rendszerezését és értelmezését, valamint a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása, elmélyítése, ellenőrzése.

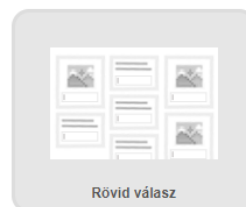
A következő feladatban a kártyán szereplő állításokról kell eldönteni és a szövegmezőbe írni, hogy igazak vagy hamisak. Az állítások hasonlóan vannak megfogalmazva, így a helyes megoldáshoz már alaposabb ismeret szükséges. Mivel a kártyák egyszerre láthatók, így a diák össze tudja hasonlítani az állításokat és ezek alapján meghozni döntését (48-49. sz. kép). A feladat célja az ismeretek elsajátításának szintjéről történő visszacsatolás.



48. sz. kép: A Btk. hatályai (állítások)
Forrás: Képernyőkép, szerzői felvétel



46. sz. kép: Tankocka elérése
A diszpozíció fajtái (Példák)
Forrás: Képernyőkép, szerzői felvétel



47. sz. kép: Rövid válasz
Forrás: Képernyőkép, szerzői felvétel



49. sz. kép: Tankocka elérése
A Btk. hatályai (állítások)
Forrás: Képernyőkép, szerzői felvétel

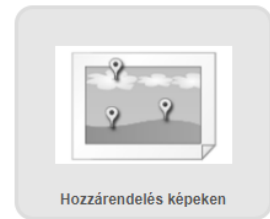
Hozzárendelés képeken

Ez a feladattípus is többféleképpen használható. A képen látható jelölőkre kattintva nagyítva jelenik meg az a rész, ahová a jelölő mutat, és alatta kis kártyákon azok az elemek, amikből ki kell választani a megfelelőt (50. sz. kép). Amit egyszer már kiválasztottak, sötét háttérrel jelenik meg.

Elérhető pedagógiai célok:

- Szemléletes, vizuálisan is elősegíti az információk rendszerezését és értelmezését, valamint a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása, elmélyítése, ellenőrzése.
- Rendszerszemlélet kialakítása vizuális információk és a közöttük lévő összefüggések alapján.

Ennél a feladattípusnál is különböző nehézségű feladatok készíthetők. Egyszerű a feladat, mikor az adott elemekhez a megnevezését kell hozzárendelni, emiatt egy elem csak egyszer szerepel, mint a következő tankockában is látható (51-52. sz. kép).



50. sz. kép:
Hozzárendelés képeken
Forrás: Képernyőkép,
szerzői felvétel

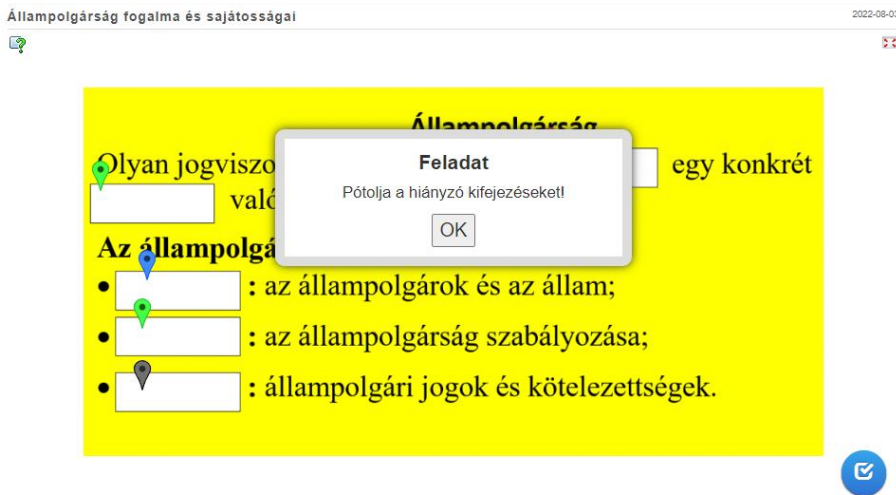


51. sz. kép: Ptk. paragrafus jelölése
Forrás: Képernyőkép, szerzői felvétel



52. sz. kép: Tankocka
elérése
Ptk. paragrafus jelölése
Forrás: Képernyőkép,
szerzői felvétel

Nehezebb a feladat, mikor a meghatározáshoz kapcsolódó fogalmat kell felismerni, kiegészíteni és/vagy a képen látható információk alapján eldönteni a csoportosítási szempontot (53-54. sz. kép).



54. sz. kép: Tankocka elérése
Állampolgárság fogalma és sajátosságai
Forrás: Képernyőkép, szerzői felvétel

53. sz. kép: Állampolgárság fogalma és sajátosságai
Forrás: Képernyőkép, szerzői felvétel

A feladattípus a kifejezetten komplex, gondolkodtató feladat készítésére is alkalmas. A következő példa jogszabály és jogforrás közötti összefüggést vizsgál, a megoldáshoz nem csupán a kapcsolódó fogalmakat kell alaposan ismernie a diáknak, hanem rendszerszemlélettel is kell rendelkeznie, a fogalmak egymáshoz való viszonyát is ismernie kell (55-56. sz. kép).



56. sz. kép: Tankocka elérése
Jogszabály vagy jogforrás?
Forrás: Képernyőkép, szerzői felvétel

55. sz. kép: Jogszabály vagy jogforrás?
Forrás: Képernyőkép, szerzői felvétel

E feladattípus is támogathatja az ismeretek gyakorlati alkalmazását: a megtanult fogalmakat példán keresztül kell alkalmazni a következő tankockában (57-58. sz. kép).



57. sz. kép: A jogi norma szerkezeti elemei
Forrás: Képernyőkép, szerzői felvétel



58. sz. kép: Tankocka elérése
A jogi norma szerkezeti elemei
Forrás: Képernyőkép, szerzői felvétel

Hozzárendelések táblázat

Ez a feladattípus a különböző elemek két vagy több megadott szempont szerinti (oszlopok és sorok) táblázatba történő rendszerezését várja el a diákoktól (59. sz. kép).

Elérhető pedagógiai célok:

- Szemléletes, vizuálisan is elősegíti az információk rendszerezését és értelmezését, a logikus gondolkodás fejlesztését.
- A tanult ismeretek begyakorlása, elmélyítése, ellenőrzése.
- Rendszerszemlélet kialakítása vizuális információk és a közöttük lévő összefüggések alapján.

Ez a feladattípus nehéz, komplex feladat. A fogalmak ismeretén túl az egyes elemek közötti összefüggések ismerete is szükséges a feladat megoldásához. A következő tankocka erre mutat példát, az általános törvényi tényállás körébe tartozó fogalmak rendszerezésével (60-61. sz. kép).



59. sz. kép: Hozzárendeléses táblázat
Forrás: Képernyőkép, szerzői felvétel



60. sz. kép: Általános törvényi tényállás rendszere
Forrás: Képernyőkép, szerzői felvétel



61. sz. kép: Tankocka elérése
Általános törvényi tényállás rendszere
Forrás: Képernyőkép, szerzői felvétel

Játékok

Az interaktív játékok, azaz a játékos feladatok a diákok belső motivációját aktiválják, amely sokkal hatékonyabb és tartósabb, mint bármely külső motiváció. A Z generációs diákok nehezen tudnak elszakadni a játékok világától, „okos” eszközeiktől, szeretik a játékos közegben meghatározó motivációs elemeket, az újdonságokat, az ismeretlen területek felfedezését, a kihívást jelentő izgalmas „kalandokat”, a versengést, a flow-élményt, azaz a tevékenységet magát. Az okostankönyvben a játékok QR kódjának kerete aranyszínű.

Mi hol van?

Ez a feladattípus a *Hozzárendelés képeken* feladat játékosított párja, ahol a tanulók egymással vagy a számítógéppel versenyezhetnek a feladat megoldása során (62. sz. kép).

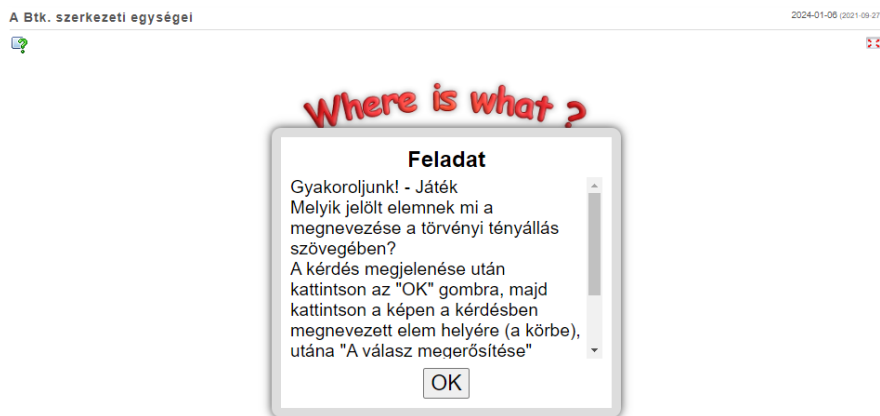
Elérhető pedagógiai célok:

- A tanult ismeretek alkalmazása begyakorlása, ellenőrzése.
- Kritikus gondolkodás fejlesztése.
- Élmény alapú tanulás, azaz a tanulási öröm fokozása.
- Motiváció növelése a szórakoztató, kihívást jelentő környezet (versenyhelyzet) biztosításával.

Példaként a 63-64. sz. képen látható feladat: „Melyik jelölt elemnek mi a megnevezése a törvényi tényállás szövegében? A kérdés megjelenése után kattintson az "OK" gombra, majd kattintson a képen a kérdésben megnevezett elem helyére (a körbe), utána "A válasz megerősítése" feliratra! Amennyiben a válasz helyes, zöld pont villan fel. Egyedüli játéknál ellenfele a számítógép lesz.”



62. sz. kép: Mi hol van?
Forrás: Képernyőkép, szerzői felvétel



63. sz. kép: Általános törvényi tényállás rendszere
Forrás: Képernyőkép, szerzői felvétel



64. sz. kép: Tankocka elérése
Általános törvényi tényállás rendszere
Forrás: Képernyőkép, szerzői felvétel

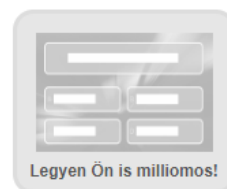
Legyen Ön is milliomos!

Ez a feladattípus tulajdonképpen a hasonló elnevezésű kvízzjáték feladatváltozata. Egyre nehezedő kérdésekre kell a diákoknak választ adni, továbblépés csak helyes válasz esetén lehetséges. Helytelen válasz esetén a játékot újra kell kezdeni (65. sz. kép).

Elérhető pedagógiai célok:

- Ismeret-alkalmazás, ismeret-ellenőrzés.
- Kritikus gondolkodás fejlesztése.
- Élmény alapú tanulás, azaz a tanulási öröm fokozása.
- Motiváció növelése a szórakoztató, kihívást jelentő környezet (versenyhelyzet) biztosításával.

Az okostankönyv fejezeteinek ismeretanyagát, mintegy összefoglalásként, fejezetenként egy-egy ilyen típusú kvíz zárja. Visszajelzések alapján megoldása szórakoztató és kihívást jelent a diákok számára. A játék 6 szintű, mindegyik szinten, mindegyik kvíz esetében 5 db kérdést tesz fel a rendszer. Egy játék alkalmával így fejezetenként 30 kérdésből szintenként egy, azaz összesen 6 kérdésre kell jól válaszolni a diákoknak a kihívás teljesítéseként (66-67. sz. kép).



65. sz. kép: Legyen Ön is milliomos!

Forrás: Képernyőkép, szerzői felvétel



66. sz. kép: Összefoglalás (5. fejezet)
Forrás: Képernyőkép, szerzői felvétel



67. sz. kép: Tankocka elérése
Összefoglalás (5. fejezet)
Forrás: Képernyőkép, szerzői felvétel

Infografika



Újfajta szemléltető eszközként az okostankönyvben megjelenik az infografika is. Az infografikát eredetileg marketing célokra alkalmazták, kevés információt hordoz, de azt látványosan. Ezen tulajdonsága az oktatásban is kiválóan használható, a lényeg kiemelésének egyik eszköze lehet. Az infografikához lehet hangot is csatolni, ilyenkor a tanuló nem csak látja, hanem hallja is a kiemelt tudáselemet. Lényege, hogy a megszerzett tudást megerősíti, ezáltal a tanuló határozottabb és magabiztosabb lesz.

A *Szándékos bűncselekmény szakaszai* c. infografika a [Canva.com](https://www.canva.com) oldalon készült, [infografika](#) (animált) és [kép](#) formában is elérhető, QR kódjának kerete türkiz színű (68-69-70. sz. kép). A Canva Edu verzió pedagógusok számára – megfelelő igazolás feltöltése után – ingyenesen elérhető.



69. sz. kép:
Infografika elérése
Forrás: Képernyőkép,
szerzői felvétel



70. sz. kép: Infografika
(kép) elérése
Forrás: Képernyőkép,
szerzői felvétel

68. sz. kép: Infografika (kép)
Forrás: Képernyőkép, szerzői
felvétel

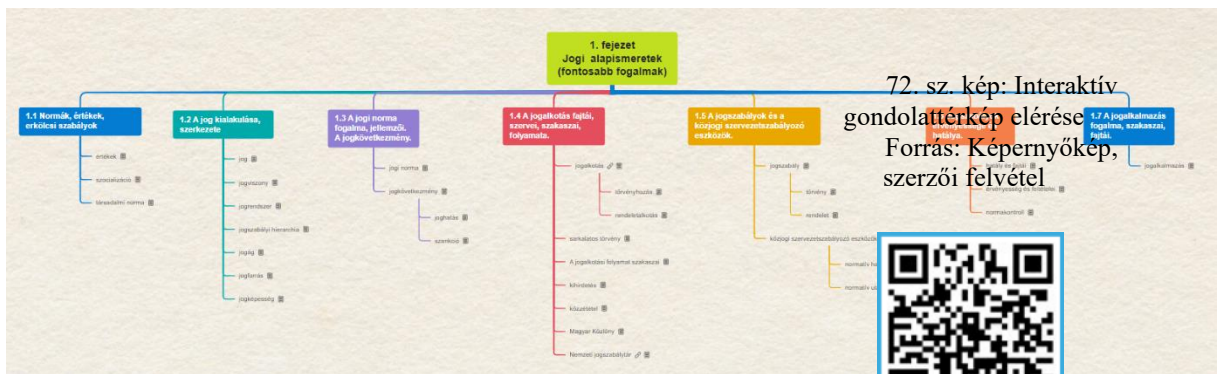
Egyéb digitális tartalmak

Minden fejezethez tartozik egy interaktív gondolattérkép, amely a fejezet fogalmait és rövid meghatározásukat tartalmazza, egy Fogalomsorsoló, amely az önellenőrzést támogatja és összefoglalásként egy-egy (korábban már bemutatott) *Legyen ön is milliomos!* játék, amelyben a kérdések egyre nehezednek és csak helyes válasz esetén lehetséges a továbblépés a következő kérdésre.

Interaktív gondolattérkép

A gondolattérkép módszer hatékonysága vitathatatlan, a tanulást és a megtanultak rendszerezését játékosítja, fenntartva az elsajátítandó anyag iránti érdeklődést. Asszociációs gondolkodásra épít, különböző elemek elhelyezésével interaktívva tehető, például az adott fogalomhoz kapcsolódó meghatározás is szerepelhet benne, így a tanuló azonnal megtekintheti, ha nem emlékszik rá. Az okostankönyvben található interaktív gondolattérképek a [GitMind.com](https://www.gitmind.com) oldalon készültek (71-72. sz. kép).

Az interaktív gondolattérképre a világoskék keret utal. Szerepe a fejezet fogalmainak tanítása és rendszerezése.



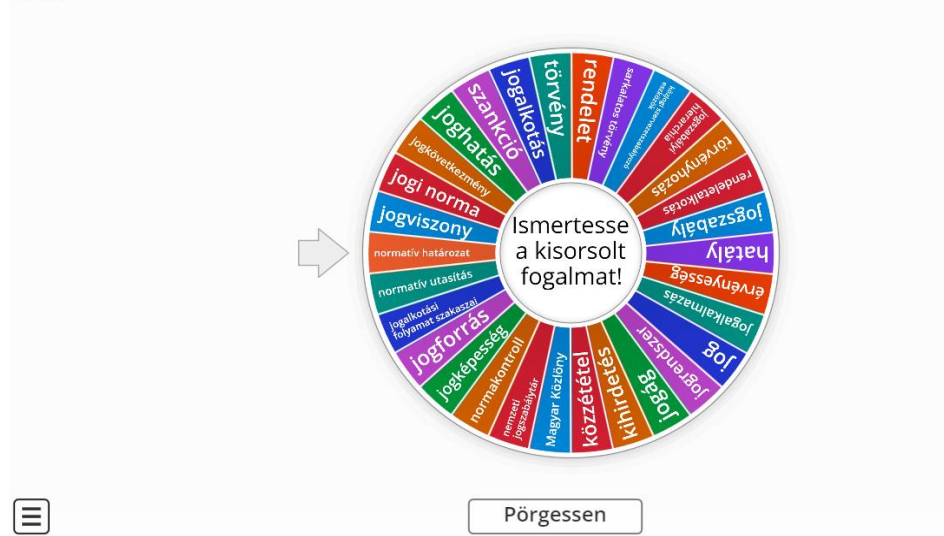
72. sz. kép: Interaktív gondolattérkép elérése
 Forrás: Képernyőkép, szerzői felvétel

71. sz. kép: Interaktív gondolattérkép (1. fejezet)
 Forrás: Képernyőkép, szerzői felvétel

Fogalomsorsoló

Az interaktív gondolattérképet jól kiegészíti a fogalomsorsoló: a diák ismerteti a kisorsolt fogalmat, majd választ ellenőrizni is tudja, így önálló tanulásra kiválóan alkalmas. A fogalomsorsolók a wordwall.net oldalon készültek, az okostankönyvben a fejezetek végén, az összefoglalásnál szürke keretű QR kóddal érhetőek el (73-74. sz. kép).

0:02



74. sz. kép:
 Fogalomsorsoló (1. fejezet) elérése
 Forrás: Képernyőkép, szerzői felvétel

Fogalomsorsoló (Jogi alapismeretek)

73. sz. kép: Fogalomsorsoló (1. fejezet)
 Forrás: Képernyőkép, szerzői felvétel

Összegzés

Magyarországon a Nemzeti Köznevelési Portál online felületén 2015-től publikusak a köznevelésben használható okostankönyvek. Mivel a belügyi-rendészeti ágazatban – tudomásom szerint – még nincs ilyen, ezért elhatároztam, mikor a célfeladatot megkaptam, hogy a hatályosítással egyidőben a segédletet digitális tartalommal egészítem ki. A célfeladat végrehajtása során felhasználtam és beépítettem az előző évben a Miskolci Rendvédelmi Technikum tanulói számára (belső használatra) készített Informatika 2. című okostankönyvem oktatásban szerzett tapasztalatait, valamint a tanulók visszajelzéseit.

A digitális feladatok alkalmazásának számos előnye van. A fentiekben is láthattuk, hogy használatukkal sokféle pedagógiai cél érhető el, emellett élményszerűbbé teszik a tanulást, fokozva a tananyag iránti érdeklődést és a diákok motivációját, ezzel is növelve az ismeretátadás hatékonyságát. A különböző nehézségű feladatok differenciálásra is alkalmasak, illetve rugalmassá teszik az oktatást, mert bármikor, bárhol elérhetők. A feladatok megoldása után a megoldás eredményéről a diák azonnal visszajelzést kap. A feladatok korlátlan számban megoldhatók, így a többszöri megoldási lehetőség során a diák fejleszti problémamegoldó képességét, analitikus gondolkodását, azaz minden ismételt próbálkozás lehetőséget nyújt fejlődésére. Lehetősége van arra, hogy javítsa korábbi hibáit és tanuljon azokból, ez hozzájárul a tanulási folyamat hatékonyságához. Az ismételt sikeres megoldások növelik a diák önbizalmát és motivációját, a sikerélmények pedig hozzájárulnak ahhoz, hogy pozitívan közelítse meg a tanulási kihívásokat.

A digitális eszközökhöz való hozzáférés kiegyenlítődni látszik, mert egyre több középiskolás kap államilag finanszírozott laptopot tanulmányaihoz, egyedül az internetkapcsolat hiánya okozhat problémát, de a legtöbb iskolában ez is biztosítva van.

Az [okostankönyv](#) a linkre kattintva érhető el.

Köszönetnyilvánítás

A szerző ezúton is köszönetet mond a Miskolci Rendvédelmi Technikum megbízott igazgatójának *Lövei László r. ezredes úrnak* és közvetlen vezetőjének *Rák László r. alezredes úrnak*, az Alapozó Ismereti Szakcsoport osztályvezetőjének támogató, biztató hozzáállásukért, az iskolai innovatív oktatási környezet (digitális táblával felszerelt tantermek, tanterem, szituációs helyszínek stb.) kialakításáért, a célfeladatban való részvétel engedélyezéséért, a feladat végrehajtásának folyamatos nyomonkövetéséért, szakmai tanácsaikért.

Mi is valójában a mesterséges intelligencia?

What is the Artificial Intelligence, actually?

DOI: [HTTPS://DOI.ORG/10.53793/RV.2024.1.2](https://doi.org/10.53793/RV.2024.1.2)

Absztrakt

A mesterséges intelligencia (MI) az emberek által végzett, nem rutinszerű tevékenységek ellátására készített számítógépes rendszerek neve, de ismertté szöveggenerálási képessége (LLM) révén vált. Meg kell értenünk alapfogalmait és működési elveit, valamint használatának következményeit energiafelhasználási, fenntarthatósági és környezetszennyezési szempontból is. Az MI lehetőségei messze állnak attól, amit feltételeznek, de veszélyei sem akkorák; feltéve, hogy megértjük, helyesen használjuk és szabályozzuk használatát.

KULCSSZAVAK: MESTERSÉGES INTELLIGENCIA, GENERATÍV, CHATGPT

Abstract

Artificial Intelligence (AI) is the name of computing systems imitating non-routine actions of humans. It became famous due to its text generation ability (Large Language Model, LLM), mainly the ChatGPT application. We need to comprehend its base notions and principles, furthermore energy consumption, sustainability, and environment damaging aspects of its operation. The possibilities of AI are far from the presumed ones, but so are also its dangers; provided that it is used and regulated correctly.

KEYWORDS: ARTIFICIAL INTELLIGENCE, GENERATIV, CHATGPT

Bevezetés

Napjaink egyik varázsszava mindannyiunk számára, függetlenül attól, hogy számítógépes feladatokkal (esetleg ezen belül mesterséges intelligenciával) foglalkozó, társ-tudományi, az alkalmazás iránt érdeklődő felelős szakember, érdeklődő kívülálló (aki vagy lelkesedik érte, vagy fél tőle, vagy csak a békés együttélés érdekében szeretne többet tudni) vagy az „utca embere” vagyunk, a mesterséges intelligencia (MI), ami az angol nyelvű kifejezés (Artificial Intelligence, AI) tükörfordítása. Az MI azért nőtt ki a tisztán tudományos tevékenység köréből, mert egyrészt akár komoly gazdasági előnyöket („újabb ipari forradalom”) is ígér, másrészt a természetes nyelvekhez idomított felhasználói felületei (a nagy nyelvi modellek (LLM) felhasználásával) tömeges felhasználást tesznek lehetővé és önálló életre keltek, harmadrészt az MI technikai kivitelezésének zárt módú kezelése magában hordozza a kiszámíthatatlan viselkedés veszélyét és az érdekcsoportok általi befolyásolás lehetőségét. A most éppen mindenhatónak és marketing szempontból abszolút szükségesnek gondolt MI-t a „vezetői összefoglalók” alapján mindegyik vállalat szükségesnek tartja versenyképességének megtartásához, a szó szoros

értelmében „bármilyen áron”. Nagy szerep jut a magukat hozzáértőnek valló, bizottságokba szervezett szakértőknek és az irányításra kényszerülő vezetőknek. A kiterjedt terület hiányos szakmai ismerete, az üzleti tőke mohósága, a társadalom és az egyének felfokozott várakozása vagy félelme, nem kellően átgondolt fejlesztésekhez és váratlan következményekhez vezet.

Összefoglaljuk a szükséges ismereteket az MI (és főleg az LLM) alkalmazása iránt érdeklődők és azt gyakorlati alkalmazások céljára felelősséggel bevezetni kívánók számára, akik nem tudnak kiigazodni a rajongók és az ellenzők szélsőséges állításai között. Az MI és az LLM felhasználói számára követhető stratégia, hogy a tanulmány általános értékelési szempontjait alaposan megértik és saját körülményeikre „benchmark” céllal (jól ismert tényekre vonatkozóan) tanulmányt készíttetnek a ChatGPT-vel, majd saját tapasztalatukból (sok munkával) döntenek el, hogy ott és úgy alkalmazható-e. Amit nagyon célszerű időnként megismételni, mivel a rendszerek folyamatos változásban vannak. Törekedtünk arra, hogy a tanulmány tartalma a kapcsolódó szakmai területekkel nem ismerős érdeklődő számára is érthető legyen

alapszinten. A speciálisabb ismereteket igénylő részek egy különálló, de elérhető Függelék elnevezésű dokumentumba (URL1) kerültek és adnak többlet információt, a fő szöveghez képest minimális átfedéssel. Jelen tanulmány főként az MI világában eligazodni kívánó és nem azt fejleszteni vagy technikai működését megérteni akaró olvasó számára szól.

Mi is az MI?

Az MI-vel kapcsolatban a tájékozódást nagyban megnehezítik az álhírek, a marketing vélemények, a valótlan tények, a valós tények elferdítése, valamint az elérhető információ (jelentős részben már az MI által manipulált) sokszínűsége. Egyik esetben sem lehet figyelmen kívül hagyni az informáltságot (a személyest, a különböző „véleményvezéreket”, a média véleményét, a hivatalból nyilatkozókat, a céges érdekeket képviselőket, a tényyszerűséget nélkülöző rajongókat és a megrettenteket), valamint az egyre nagyobb keresletet kihasználó, egyre növekvő számú (és sokszor elszabadult fantáziájú) „szakértőt” is beleértve. Arra, hogy hogyan is működik az MI, a magukat szakértőknek tartó művelői és felhasználói – sőt fejlesztői – is egyre kevésbé tudnak válaszolni; szélsőséges értékelésű hatásai pedig kiterjednek az oktatástól a művészetekig, a gazdaságtól a tudományig; felelőtlen fejlesztői által könnyen elérhetővé tett fejlesztőanyagai következtében egyre több területen jelennek meg és befolyásolják életünket (Kovács 2023).

A tanulmány *nem* MI-ellenes, hanem MI-kritikus. Igyekszik feltárni a területtel kapcsolatos felfokozott technológiai várakozással (valamint éppúgy a teljes kiábrándulással) szemben álló másik oldal érveit is. Megpróbálja – a teljesség igénye nélkül – a nagyon sokszor nem egyértelmű alapfogalmakat és összefüggéseket definiálni, azokat egységes és viszonylag átlátható keretbe helyezni, egymáshoz való viszonyukat tisztázni, továbbá az MI-t csupán alkalmazni kívánó szakember számára érthetően megfogalmazni a társ-tudományainak állításait. A tanulmány rövid és koherens módon tárgyalja a szükséges alapfogalmakat és vezeti be az olvasót az elemeiben ismert fogalmak összefüggő rendszerébe, valamint további tájékozódáshoz ad útmutatást az érdeklődő olvasó számára.

Az MI meghatározása

A természetes intelligenciát (TI) általában az emberi létezéshez (vagy legalább fejlett biológiai rendszerhez kötött (Macpherson et al. 2021), biológiai szövetek bonyolult együttműködésén alapuló magasabb rendű viselkedési módnak tekintik, amelynek viselkedését sem a legmagasabb agyi funkciók, sem az elemi neurális

műveletek területén nem ismerjük teljes mértékben (Buzsáki 2019). Úgy tekintjük azonban, hogy a biológiai tanulás (az intelligens működés) lényegi része, hogy a rendszer módosítani tudja saját jövőbeli működését, amihez „jövőbe látás”, azaz „intuíció is szükséges”. Lényeges momentum, hogy nem csak egyszerű mennyiségi tanulásként, inkrementálisan, a korábban már látott viselkedésminták valamelyikét (esetleg azok valamilyen kombinációját) használjuk, hanem egy új és váratlan helyzetre is jól reagáljunk (nem-inkrementális, ugrás-szerű tanulás). Augusto és Gambino (2019) meghatározása szerint *a kizárólag a korábban látott viselkedésmintákon alapuló rendszer nem intelligens: nem szerez új képességet, hanem változatos módokon próbálkozik*. Ennek alapján meg lehet különböztetni *betanulást*, azaz viselkedésminták elsajátítását, ami esetleg több megtanult viselkedésminta valamilyen módon való kombinálását is jelentheti, de a viselkedés csak a már látott mintákra korlátozódik; és *megtanulást*, ami az addigi képességekhez hozzáadja az új helyzetben szükséges viselkedés képességét (Végh 2024). Lényegében a betanított munkás és szakmunkás, ill. az asszisztens és menedzser/orvos/tanár megkülönböztetésével analóg: másfajta tudásról és másfajta képességről van szó. Más szavakkal kifejezve, ezzel teljesen összhangban van az a meghatározás, amivel a ChatGPT (a fejlesztői által összeállított tudásbázis alapján) megkülönbözteti a gépi (és az ennek alapjául szolgáló mesterséges) és az emberi intelligenciát. Mindegyikre szükség van a megfelelő helyen és összefüggésben. Az MI matematikai vonatkozásainak (és a fogalomnak) jó megfogalmazását tartalmazza a *Mi az MI?* című könyv (URL2), annak ellenére, hogy negyedszázada íródott.

A kiegészítő dokumentumban bevezetünk további, (kényszerűségből) jelzővel ellátott intelligencia fogalmakat is, amik részben a biológiai és technikai rendszerek különbözőségéből fakadnak, részben a biológia működés félreértése miatt használt probabilisztikus, valamint a technikai megoldások következményeként előálló kvázi-probabilisztikus működésre utalnak. A digitális számítógép használatának mind tömegesebb elterjedésével előállt lehetőség egyre több terület ad az emberi intellektus előállítását célzó próbálkozások számára, bár mint Chu–Prokopenko–Ray (2018) magyarázza: „*szigorúan véve, ma sem értjük, milyen elv alapján számol az élő anyag*”. Úgy pedig igen nehéz valamit mesterségesen előállítani, hogy nem tudjuk, az eredeti hogyan működik és – az átgondolatlan technikai megvalósítások miatt – egyre kevésbé tudjuk ellenőrizni a megvalósított másolatot, viszont különféle érdekektől motiválva az internet segítségével tömegesen és korlátozás nélkül használjuk.

Az MI, mint tudományos terület

A számítógépet és az agyat tanulmányozó tudományok egymást nem mindig előnyösen befolyásoló kölcsönhatásban állnak egymással (Macpherson et al. 2021; Hassabis et al. 2017). Mindkét területen használják a társ-tudományok, például a matematika, a fizika, az informatika, az elektronika, a számítógéptudomány, az idegtudományok eredményeit. Mindegyik terület képviselői saját területük elsőrangú szakemberei és ezt feltételezik a vonatkozó társ-tudományok képviselőiről is, de magára a társ-területre vonatkozó tudásuk erősen korlátozott. Minden tudományterület valamilyen közelítéseket használ a természet jelenségeinek leírására és ezeknek a közelítéseknek (sokszor csak implicit) érvényességi köre van, ami már sokszor a társ-tudományok képviselőinek figyelmén kívül marad, így esetlegesen a társ-tudományok eredményeit és következtetéseit olyan területekre is alkalmazni akarják, ahol a közelítések már nem alkalmazhatók. Jellemző példái éppen a számítógéptudomány és a bioinformatika. Az MI ezeknek a társ-tudományoknak egymásra hatásából született; mára (legalábbis méreteit és kutatási intenzitását tekintve) lényegében önálló tudományterületté vált és magán viseli az említett „túl-használat” jellemzőit.

Fogalmi tisztázás

A legtöbben már ott elakadnak, amikor magát az MI fogalmat próbálják megérteni, így a kapcsolódó fogalmak megértése reménytelené válik. Tekintve, hogy már a TI megfogalmazásakor is bizonytalankodtunk, nem is csoda: azt sem tudjuk pontosan megfogalmazni, hogy mihez, és annak melyik részéhez hasonlítjuk. A bizonytalan megfogalmazású TI és a marketing ihlette MI megfeleltetése egymásnak komoly értelmezési nehézségeket okoz (Kar–Kornblith–Fedorenko 2022). 2023-ban talán a legismertebb MI-vonatkozású megvalósítás, a ChatGPT álláspontja (azaz a rendszer tudásbázisát készítő OpenAI fejlesztők által összeállított szöveggészlet tartalma) ellentmondásos. Egyrészt korrektil fogalmaz önmagáról: „*Én egy gépi intelligencia vagyok, amely mesterséges intelligencia alapokon működik*”, azaz egy lényegében tisztán matematikai módszer műszaki megvalósítása. Másrészt viszont, ha technikai rendszerként általa okozott károkról (energiafelhasználás és környezetszennyezés) kérdezik, akkor megtagadja és csupán egy szoftvernek minősíti önmagát, azaz csupán a szoftvert futtató hardver okoz kárt; miért használnak hardvert.

Az MI fogalma

Az MI Fang–Su–Xiao (URL₃) által adott definíciója az emberi intelligencia közelebbéről nem definiált fogalmához vezet vissza: „Az MI abban különbözik más berendezésektől, amelyeket emberi tevékenységek javítására vagy helyettesítésére terveztek, hogy *nem csak rutin vagy ismétlődő tevékenységek megoldásán alapszik, hanem az intelligencia-szintű emberi viselkedésen is.*” Ez utóbbi meghatározás – lényegében helyesen – kizárja fogalomköréből az egyszerű (főként visszacsatoláson alapuló) szabályozó és vezérlő köröket, az egyszerű osztályozási és mintafelismerési eseteket, a bonyolult döntési fák és optimalizálási eseteket, a statisztikai számítások és mintázatok keresése alapján hozott döntéseket, az idősorok vizsgálata alapján megalapozott előrejelzések készítését stb., viszont fontosnak tartja az emberi viselkedés imitálását. Megfontolandó lenne Augusto és Gambino (2019) meghatározását használni az MI-re is: „az intelligencia egyrészt lehetővé teszi külső ingerekre megfelelő válasz adását, másrészt *további képességek megszerzését*”.

Már csak az emberi tevékenységek sokrétősége is megnehezíti azok imitálásának csoportba sorolását. A megvalósítások technikai és matematikai háttere, valamint társadalmi/technológiai hatásai és népszerűsége/elterjedtsége miatt *az MI-nek lényegében nincs egységes fogalma és kezelése.* Ehhez még hozzájárul a megvalósításban és felhasználásban (főként gazdaságilag és politikailag) érdekelt marketingje, különböző hiedelmek és várakozások terjesztésében nagyon, tényszerű ismeretek megszerzésében sokkal kevésbé aktív egyének és csoportok tevékenysége, továbbá a fejlesztés gazdasági hasznából részesedni kívánók felelőtlen magatartása és az „érdekes, szórakoztató és hasznos” technológiát „ingyen” használni kívánók érdeklődése. Ennek megfelelően maga a fogalom is gyorsan változik, és a változáshoz a szakmai-technológiai háttérnek van a legkevesebb köze. A technológiai várakozás gyors felfutása idején kiadott 2018-as Európai Unió (EU) meghatározás még főként az automatizálást (elsősorban érzékelés, egyszerű számítások és beavatkozások; osztályozás; matematikai algoritmusok gyors végrehajtása) értette alatta, de jövő idejű megfogalmazással, a hívők erősen csoda-váró hangulatában irreális távlati célokat tűzött ki. Ráadásul (a marketing ihletésű vezetői összefoglalók szintjén) összekavarodnak a csoda-váró rajongói vágyálmok, az ide vonatkozó matematikai módszerek használata és fejlesztése az ezeket megvalósító eszközök használatával. A fogalmat 2022-ben az EU az eredeti jelentéstől teljesen elrugaszkodva, gyakorlatilag a ChatGPT csevegőprogram testére szabta: (tanításhoz) *bemenő adatokat használó és generált tartalom kimenetet előállító, egy bizonyos listában szereplő technológiákkal*

előállított szoftverként határozta meg (ennek alapján a négy év előtti megfogalmazás, az automatizálás, már nem tartozik az MI fogalmába; biztosan nem felel meg az MI eredeti modellezési szándékának sem; továbbá erősen kérdéses, hogy maga az antropomorf ambíciójú MI belefér-e a testére szabott kategóriába). Az európai álláspont nyomán született 2019-es Magyarország Mesterséges Intelligencia Stratégiája (URL4) is „fizikai eszközigény nélkül” képzelel el az MI fejlesztését és használatát (és természetesen nem látta előre az említett fogalomváltozást).

Az álláspontok szinte egyetlen közös vonása, hogy a legelső számítógépes korszak túlzó és nagyjából megalapozottság nélküli, időközben már többszörösen csalódáshoz vezető elvárásai alapján tűzik ki céljaikat. Egyszerűen nem akarják tudomásul venni a valóságot tartalmazó szakértői jelentéseket: *„Hat évtizeddel később a kitűzött kognitív képességek még mindig elérhetetlennek tűnnek”* (Jordan 2019). Kétségtelen tény, hogy az évtizedekkel ezelőtti állapothoz képest a matematikai módszerek, továbbá a számítási technológia és erőforrások használata óriási mértékben fejlődött, de ettől a kitűzött cél megalapozottság nélküli vágyalom maradt. Maga a terület azonban a több rendelkezésre álló hardver és adat alapján ambiciózusabb és olyan elvárásokat támaszt, mint legutóbb a „dot.com” buborékkal kapcsolatban; hasonló eredményesség várható. Ezt a várakozást az sem befolyásolja, hogy a manapság legnépszerűbb csevegő alkalmazások (készítői által írott adatkészlet) tisztán „nem”-mel válaszolnak arra a kérdésre, hogy tudnak-e gondolkodni, önálló felfedezéseket tenni, dolgokat/eseményeket előre jelezni, a természetes intelligencia számára megoldhatatlannak bizonyult feladatokat megoldani. A rajongók irreális fantázia-álmái önállóvá váltak és nem hagyják befolyásolni magukat a keserű tények által. Ebben a képben nem jut hely sem a technológia tömeges felhasználásával okozott károknak, környezeti ártalmaknak, energiafelhasználásnak, sem a tömeges társadalmi felhasználás által okozott károknak és veszélyeknek. *Figyelemre méltó, hogy a stratégiákat összeállító és a technológia használatát szabályozni kívánó „szakértők”, politikusok és jogászok milyen kevéssé ismerik a technológia elméleti és gyakorlati tulajdonságait, továbbá használatuk következményeit. Megelégszenek az érdekelt cégek által készített, a vezetőknek szánt hangos marketing anyagok ismeretével. Még az sem tűnik fel, hogy – az utólagos felelősségre vonástól félve – a fejlesztő cégek egyre inkább követelik a terület fejlesztésének és használatának szabályozását, azaz saját önállóságuk csorbítását.*

Az MI megítélése

Ez a bizonytalanság nem véletlen. Az EU által elfogadott jelenlegi (2023) MI szabályozás (URL5) elismeri, hogy *nincs általánosan elfogadott meghatározás (jellemző módon gépi intelligenciáról beszél, de MI-t mond)*. Hivatkozás nélkül, de egyetért Lighthill (1972) (URL7) fél évszázados kutatási jelentés megállapításával, hogy *fogalmilag sem lehet egységes MI-ről beszélni. Nem csak működési módját és tulajdonságait nem ismerjük, de még a fogalmát sem tudjuk pontosan meghatározni. Bevezetéséről (megvalósíthatóságáról és hatásairól) sem készült tanulmány. Részleges tiltásának dokumentumában (URL6) is csak az szerepel, hogy a (szakmai) „előadó a mesterséges intelligencia-rendszer nemzetközileg elismert fogalom meghatározásának használatát szorgalmazza”* (figyelmen kívül hagyva, hogy pár sorral odébb állította, hogy nincs olyan).

A betiltás és a hiányzó fogalmi meghatározás mellett legalábbis furcsán hat, hogy *„A mesterséges intelligencia egy kiforrott és használatra kész technológia, amely felhasználható az ipari folyamatok során keletkező, egyre növekvő adatmennyiség feldolgozására.”* (Ami a vezetői összefoglalóban található marketing szöveg; magának az MI-nek, pontosabban az azt belülről ismerő fejlesztőknek más a véleménye: *„egyáltalán nem tükrözi a jelenlegi állapotot a mesterséges intelligencia területén”*. Ez az állítás sikeresen összemossa a vágyalmokat, a különböző matematikai módszereket megvalósító rendszereket, a különböző technikákat, a különböző alkalmazási területeket; a néhány valóban sikeres módszer, technika és terület álcája mögé rejtve a valójában fogalmilag sem ide tartozókat.) Erre vonatkozóan a magyar stratégia (URL4) helyesen állapítja meg, hogy az *„általános MI kutatása fejletlen és bizonytalan”* (de az akkor is igaz marad, ha használatát leszűkítjük egy-egy területre). *Két különböző intelligenciáról van szó: a gépi intelligenciát kutatjuk és az emberi intelligenciát próbáljuk elérni*, lásd még az MI véleményét is a Függelékben (URL1). *Csupán közelítőleg lehet az MI fogalmait a TI hasonló fogalmaira leképezni* (Kar–Kornblith–Fedorenko 2022). A Jordan (2019) által javasolt sorrendfordítás (AI helyett IA, Intelligence Augmentation) egyáltalán nem csupán szójáték. A számítógépes rendszer óriási ismeretanyaghoz fér hozzá, több adatot tud tárolni, gyorsabban működik, mint az ember; jól kiegészíti és segíti az emberi intelligenciát (természetesen feltéve, hogy helyesen valósították meg), de nem pótolja azt. Rutin feladatokra és a leggyakrabban előforduló esetekben akár teljes értékű megoldásként használható, de mindenképpen a TI felügyelete alatt.

A működési mód teljes félreértésének (vagy inkább az erre vonatkozó tudatlanságnak) a következménye, hogy a félrevezetett társadalom elitnek mondott rétege olyan követelményeket támaszt az MI-vel kapcsolatban, mint átláthatóság, kiszámíthatóság, elszámoltathatóság, korrektség, bizonyos elvárások és értékek tiszteletben tartása. Amire egy tapasztalati eloszlásokon és matematikai valószínűségeken alapuló rendszer még bonyolult (és TI alapú) „MI algoritmusok” használatával sem képes, és amit csak tetéz a megismerhetőségre (adatkészletek, algoritmusok és technikai eszközök) vonatkozó, céges érdekek védelmére szolgáló korlátozás.

Bár a 70-es években, a már kicsit haladottabb számítógéptudomány korszakában (a fizikai Nobel-díjas) R. P. Feynman (2018) felhívta a figyelmet arra, hogy jobb lenne a „haladó algoritmusok” (advanced algorithms) kifejezést használni a területre. Kicsit részletesebben talán „az ember működésének valamely aspektusához hasonló viselkedést legalább közelítőleg imitálni próbáló matematikai algoritmusok” lenne rá egy jó összefoglaló név. Ezeket az elméleti módszereket pedig csak konkrét (matematikai, informatikai, elektronikai és újabban biológiai) megoldások használatával lehetséges megvalósítani, tulajdonságaik megértéséhez pedig a használt technológia tulajdonságait kell figyelembe venni. Ezen megoldások jelentős korlátozásokat szabnak mind a megvalósítás, mind a felhasználás számára; sőt, a megvalósítás döntő módon befolyásolja a tapasztalható tulajdonságokat.

Az elnevezés helytelen használata annyira elterjedt, hogy a robbanásszerű felfutás idején vezető elméleti tudós is kifakadt ellene: „*Ne hívjunk mindent MI-nek*” továbbá „*valójában miért nem intelligensek a mai rendszerek*” (Pretz 2021). „*A rosszul megválasztott elnevezés megakadályozza a vonatkozó intellektuális és felhasználási problémák megértését*” (Jordan 2019). A kapcsolat fordítva is igaz: az MI módszereinek biológiai relevanciája legalábbis kérdéses (Macpherson 2021). Mindennek ellenére, az elterjedésnek immár hagyományteremtő ereje van, pedig a valódi ok, amint egy újságíró a Google fejlesztői konferenciájával kapcsolatban megfogalmazta (Jordan 2019): „*az MI gyorsan olyan marketing kifejezéssé vált, ami csak annyit jelent 'számítógépek által végzett számítógépes tevékenység'*”.

Adatfeldolgozás MI módra

Technikai kivitelüket illetően, a rendszerek jellemzően elektronikus eszközök (változó mennyiségű és nem mindig helyesen értelmezett biológiai imitációval). A jellemző eszköz a számítógép, illetve az azon futó program; de vannak speciális (számítógép-szerű, jellemzően nagyrészt analóg) cél-berendezések is. Amint a Függelék tárgyalja (URL1), a számítási folyamatokat lehet vezérelni adatokkal, amelyek

megkeresik a feldolgozásukhoz szükséges utasításokat; közvetlenül utasításokat adhatunk egy-egy adat(csoport) feldolgozására, vagy egy esemény hozza működésbe az előbbieket valamelyikét.

Biológiai lényként az ember is adatvezérelten működik (legalábbis olyan értelemben nincs vezérlőegysége, mint a számítógépnek). Például a szemünkkel látott élmények olyan adatokat generálnak, amelyek bekerülnek neuronjaink hálózatába, ott a rendkívüli sűrűséggel összekapcsolt neuronok hálózatán egy jól meghatározott pályán végig haladva átalakulnak, és cselekvésben vagy emlékezésben nyilvánulnak meg. Lényegében ezt a funkcionalitást szeretnénk megvalósítani technikai elemek (jellemzően számítógép) felhasználásával: pl. a robotok adatból (vezérlés vagy érzékelés) cselekvést valósítanak meg. A mesterséges neurális hálózatok lényegében logikailag ezt teszik, technikai megvalósításuk (ezért működésük és tulajdonságaik is) azonban lényegesen különbözik.

A számítógépek elemi programutasítások sorozatának végrehajtásán alapulnak; egy-egy bonyolultabb feladat megoldására számos elemi utasítást végrehajtó programot kell írni és annak futtatása eredményezi a megoldást. Egy ilyen lehetséges program annak imitálása, hogy a megkapott adatokból – lényegében eléggé általánosan megfogalmazott feladatot végrehajtó program (egy algoritmus) felhasználásával – látszólag programozás nélkül állítjuk elő a kívánt eredményt. Az ilyen (általunk ál-MI-nek nevezett) megoldásban ugyan az adatok vezérlik a feldolgozást, és az „MI algoritmus” módosításán keresztül (az általános megfogalmazást bizonyos kérdések esetében speciálissá téve) módosítani tudjuk a számítógépes programot, de az minden módosítástól bonyolultabbá válik és működése egyre közelebb kerül ahhoz, amit közvetlen programírással (TI alapján) el tudnánk érni. (Vegyük észre, hogy három szinten is elrejtjük a TI-t. Az adatkészletet TI készíti (a nyers szövegeket és a szövegösszefüggéseket is); az algoritmusokat TI írja; az újabb keletű, ún. „prompt engineering” szakma pedig immár nyíltan a TI beavatkozása: ha még az előző két elemmel nem működik kielégítően, akkor a paraméterek „finomhangolásával” a TI beavatkozik, hogy (arra az egy kérdésre) minél tökéletesebb válasz szülessen.) Az algoritmus-vezérelt működésnek mind a megvalósítást, mind a működést, mind annak eredményességét illetően komoly korlátai vannak. Az egyik alapvető probléma az ún. skálázhatóság: a felületes hasonlóság alapján készített („játék”-szintű) modell egyáltalán nem működik valós méretű problémák megoldásakor.

Turing-teszt

Alan Turing javasolta az intelligencia elfogadhatóság mérésére azt a tesztet (bár inkább odavetett ötletnek, mint rendszeresen használható tesztnek szánta), hogy a mesterséges rendszer akkor mondható intelligensnek, ha egy ember egy rövid társalgás során nem tudja eldönteni, hogy emberrel vagy géppel lépett kapcsolatba. Azaz, a Turing által eléggé nagyvonalúan megfogalmazott emberi szereplő intelligenciájától, a kommunikáció eszközeitől, a feltett kérdések körétől, az elérhető adatkészletek körétől és időszerűségétől, a csevegés időtartamától, a megengedett „gondolkodási” időtől stb. is függ a teljesítés. A kérdésfeltevés lényegében értelmetlen, lévén, hogy két teljesen eltérő, csupán nevében egyező intelligenciát hasonlítunk össze. A Függelék szerint (URL1) az ún. generatív eloszlásfüggvénnyel dolgozó rendszerek definíció-szerűen átmennének, viszont az emberek többsége nem feltétlenül menne át a Turing-teszten.

A Turing-teszt sikeres teljesítésének demonstrálására irányuló lobbizás igen erőteljes. Az érdeklődést felkeltő publikus „*A ChatGPT teljesítette a Turing-tesztet*” (Biever 2023) (ezzel szemben áll a ChatGPT önértékelése, lásd a Függelékben (URL1)) szalagcím alapján a már csak korlátozottan elérhető közlemény címe „*Könnyű intelligencia tesztek, amelyeket a ChatGPT nem tud teljesíteni*”, a közleményben pedig az áll: „*a GPT-4 és más alapú rendszerek valószínűleg átmennének a népszerű Turing-teszten olyan értelemben, hogy az emberek nagy részét bolonddá teszik, legalábbis egy rövid csevegés során*”. A szöveggenerálásra elég jó hasonlat: egy jó minőségű márkás árucikk helyett gyártunk olyan utánezetet, amelyik rövid időn keresztül meg tudja tévesztetni azokat az embereket, akik nem ismerik az eredetit. Aminek természetesen csak akkor van értelme, ha el tudjuk adni eredetiként (a párhuzam a ChatGPT felhasználásával nem véletlen). Munkavállalók MI-vel való helyettesítésére is igaz.

Az értékelés szerint az MI alapú rendszerek nem felülmúlják az embereket, hanem a használt benchmark képességei végesek (és a rendszereket „felkészítik” a tesztek teljesítésére, lásd a $2+2=?$ feladatot). A „contamination” jelensége (előbb-utóbb felbukkannak a helyes válaszok is a betanításhoz használt adatkészletben) révén a felhasznált nagy mennyiségű adat között megtalálja a helyes választ. *Feltéve, hogy az adatkészletben megtalálható és ott csak helyes válaszok szerepelnek.* Ha nem nyert, „újra húzhat”. Ha nem sikerült, az nem újsághír. Aztán a fejlesztők tökéletesítik a rendszert, hogy azt az egy benchmarkot teljesítse.

Az MI története

Az MI szinte az emberiséggel egyidős és kötődik az emberi fantáziához. Az ember nem csak fantázialényeket teremtett, hanem valódi (ember alkotta) teremtmények működését is antropomorfizálta: saját (megtapasztalt és többé-kevésbé megértett) működéséhez hasonlította. Hogy manapság számítógép és elektronika a nyersanyag, az már az első MI „forradalom” idején is magától értetődött. A második idején nyilvánvalóvá vált, hogy a „haladó algoritmusok” nem elegendőek („expert knowledge” és megfelelő interfész is kell), a mostani harmadik „aranykort” pedig elsősorban az elképesztően nagy mennyiségű adat felhasználása motiválta és dominálja; széles körű használatát a számítógépes nyelvészet valóban imponáló fejlődése tette lehetővé. Vegyük észre az alábbiakban: minden eddigi megvalósításban csupán arról van szó, hogy valamilyen trükkkel elbújtatjuk a TI eredményeit, és megpróbáljuk a gyanútlan felhasználóval elhithetni, hogy az általa látott eredmény az MI terméke. Végül is, *az emberi intelligencia mesterséges megvalósítása azt jelenti, hogy a mesterséges szerkezetnek antropomorf viselkedést és kinézetet adunk azzal a céllal, hogy a többi embert megtévesztjük.*

Nem-számítógépes MI

Az EU (URL14) 2018-ban így definiálta az MI-t: „*A mesterséges intelligencia intelligens viselkedésre utaló rendszereket takar, amelyek konkrét célok eléréséhez elemzik környezetüket és – bizonyos mértékű autonómiával – intézkedéseket hajtanak végre*”. Eme meghatározás szerint sok száz év óta, naponta használt rendszer például a vízszint szabályzó, aminek konkrét célja a vízszint állandó értéken tartása és a környezet folyamatos elemzése során a vízszint csökkenésének észlelésekor autonóm módon intézkedik: kinyit egy szelepet és vizet enged a tartályba, majd a vízszint helyreállítása után autonóm módon elzárja a szelepet. A TI a szelepebe van építve. Az ókori vízórák „vezérlőszervezete” egy felelős rabszolga vagy pap volt (beépített TI).

Több szempontból is jó példa az MI-vel kapcsolatban az a mechanikus szerkezet, amelyiket a zseniális mérnök, Kempelen Farkas 1769-ben mutatott be „sakkozógépként”. Sokkal bonyolultabb tevékenységet (de lényegében továbbra is egyetlen célt): gépi megvalósítású sakkpartnert hozott létre. Lényegében az emberi test és az emberi agy részleges modellezése volt a cél – mechanikusan, és valójában a TI bevonásával. Az antropomorf jelleget sakkfigurák mozgatása, a bábu fejének és kezének mozgatása stb. valósították meg; a lényegét pedig, a tudásbázist a TI szolgáltatta: szó szerint elbújtatta a TI-t a látszólagosan

MI megoldásban. Abban a korban már bonyolult óraművek (és általuk mozgatott figurák) léteztek, de nem lehetett velük kommunikálni (nem voltak interaktívak); lényegében ezt az igényt elégítette ki a sakkozógép.

Mai terminológiával úgy mondhatnánk, hogy a sakkozó asztal és a török figura alkották a felhasználói interfészt, a működtető mechanika az igénybe vett infrastruktúrát, a tudásbázist (avagy adatkészletet) pedig a TI (emberi sakkmester) szolgáltatta, közvetlenül, valós időben. Erre sem igaz, hogy az MI legyőzte volna a TI-t: egy jól sakkozó ember győzte le a neves embert, illúziót keltően használva a bonyolult infrastruktúrát. A sakkozó automata új elemként (igaz, nagyon korlátozott módon) *kommunikál* az emberi felhasználóval (Kempelen Farkas egyéb tevékenységei – pl. a beszélőgép alapján – feltételezhetjük, hogy ez volt a valódi érdeklődése). Még akár azt is mondhatjuk, hogy ilyen értelemben az általa készített automata átment a Turing-teszten: az emberek túlnyomó többségét megtévesztette. Itt már jól látható az embernek az az igénye, hogy elfogadja a mesterséges teremtmény korlátozott valóságát; azzal együtt, hogy látja annak „bábu” jellegét, szögletes mozdulatait, a vizuális és akusztikus kommunikáció hiányát stb. Tudomásul veszi, hogy az illúzióknak korlátai vannak.

Illúziót kelteni természetesen szavakkal és képekkel is lehet; feltéve, hogy a felhasználónak szándékában áll elhinni az illúziót; lényegében ezen alapszik a mesék és legendák többsége. Felhasználóinak nagy része számára az MI alapú termékek használata olyan érzést biztosít, mint az „Aliz kalandjai Csodaországban” (Alice's Adventures in Wonderland, 1862) közismert satirikus fantázia-mese vagy az „Óz, a csodák csodája” (The Wizard of Oz, 1939) főhősenek az az élmény, hogy egy olyan fantázia-világba kerül, ahol *az állatok és a fantázia-teremtmények antropomorfizáltak*, továbbá a környezet és az események a valós és a szürreális keveréke. Vegyük észre, hogy ezek a teremtmények kommunikálnak (azaz beszélnek és viselkednek), továbbá – bár furcsán – gondolkodnak.

Egy másik irodalmi párhuzam lehet a híres regény „Don Quijote de la Mancha” két világa: az elképzelt lovagi és a valóságos világ. Ahol szintén keverednek a valós és szürreális elemek: az ábrándképben szereplő óriások a valós világban szélmalomok és az ábrándképben rohamozó hős lovag valójában egy ábrándkép által bolonddá tett, jó szándékú, naiv ember. Itt a társadalom bolonddá tett része játssza Don Quijote szerepét és az MI-ból az „intelligencia” szóhoz kötődő ábrándkép a lovagregények légkörét. Az ember, a szélmalom és a rohamozás is valóságos, csak szerepük látszik másnak a két világban. Saját világában mindegyik konzisztens; a baj (ott is) akkor következik be, amikor a képzelte és a valóságos világ keveredik (komoly célra, munka-

segítőként használjuk). Jelen tanulmány szerzőjének jut Sancho Pansa szerepe: igyekszik a szegény, bolonddá tett társadalmat a valóság talajára visszarángatni.

Számítógépes MI

A ChatGPT-t létrehozó OpenAI világosan megkülönbözteti a kizárólagosan elméleti erőforrásokat használó „mesterséges intelligencia” és a technikai erőforrásokat használó „gépi intelligencia” fogalmakat. Mai felhasználói, stratégiakidolgozói, szakértői stb. viszont nem tudnak ilyen megkülönböztetésről. Ezért is fontos lenne a fogalmak tisztázása. A két fogalom összetévesztéséből származik a magyar MI stratégiában (URL4) is tetten érhető nagyon naiv álláspont, miszerint *„A mesterséges intelligencia egy globális technológia, határok és fizikai eszközigény nélkül”,* valamint a – különösen a globális energiaválságban szenvedő korunkban – fájó tévhit, hogy *„Ezúttal azonban nincs a fejlődésnek természeti erőforrás-igénye”.* Aminek maga a Stratégia szövege is számos ponton ellentmond: *„Cél a kutatási és általános fejlesztési feladatok elvégzéséhez szükséges erőforrások központi infrastruktúrával való kiszolgálása, a létrejövő infrastruktúra igény szerinti folyamatos fejlesztése, hosszú távú fenntartása.”* Amennyiben csupán elméletileg gondolkozunk különböző módszerekről, akkor valóban csak gondolkodó fők által kifejlesztett matematikai módszerekről kell beszélnünk; ez az MI. Ha azonban azokat technikai eszközökkel (jellemzően számítógépes programokkal) valósítjuk meg, akkor a képbe bevonul a technikai megvalósítás a természeti törvényekből következő korlátaival együtt, ami a *gépi intelligencia*.

Nem lehet az óvatossági és hihetőségi szempontokat sem figyelembe venni, valamint hibákat és sebezhetőségeket sem megérteni az elméleti és technológiai háttér tárgyalása nélkül; különösen, mivel a technológiai fejlődés iránya egyre kiszámíthatatlanabb és kockázatosabb. Figyelembe kell venni energetikai és környezetszennyezési szempontokat, mert az EU 2023-as álláspontja szerint (URL6) is: *„A mesterséges intelligencia-technológiák és adatközpontok nagy szénlábnyommal rendelkeznek a megnövekedett számítási energiafogyasztás, valamint a tárolt adatok mennyisége és a keletkezett hő-, elektromos és elektronikus hulladék mennyisége által előidézett magas energiaköltségek miatt, így még nagyobb lesz a szennyezés.”* Pedig az már az MI pár év előtti felfutása előtt is magas volt (Jones 2018).

Sok elméleti tanulmány lehetőségként említi, hogy *az MI csökkentheti bizonyos tevékenységek lábnyomát, annak említése nélkül, hogy az ennek elérésére szolgáló technológia használata mennyivel többet növel rajta.* Egyetlen, az újabb keletű gigantomániás szellemben épített rendszer betanítása több tíz tonna szén-dioxid kibocsátásával egyenértékű (URL8), a GPT-3 több mint ötszázal. Az ilyen rendszerek betanításának

energiaigénye gigawatt-óra nagyságú és hosszú hetekig/hónapokig tart (Luccioni, Viguier és Ligozat 2022-es cikkükben 118 napot említenek) (URL8). Nem elegendők a merész célkitűzések és ambíciók; a technológia pontos megértésére és megvalósítására van szükség, a csodavárás nem stratégia. Esetleg a kidolgozók figyelmét fel lehetne hívni pl. arra, hogy a közgazdaságtan szerint a haszon eléréséhez befektetés és termelési költség is szükséges; vagy legalább arra, hogy ingyen ebéd nincs.

A technikai és biológiai számítási módok alapvetően eltérnek; a biológiai működés számítógépes imitálásának elve önmagában több nagyságrendi számítási határfokcsökkenést okoz (Végh 2021), ami a teljesítmény fokozására használt megoldások következtében mára odáig vezetett, hogy az igazán fontos problémák (pl. globális felmelegedés, személyre szabott genetikai gyógyítás (Conklin–Kumar 2023)) megoldásának lehetővé tételéhez a számítási határfoknak kb. 11 nagyságrendet kellene javulnia. Az egyértelműen az MI által gerjesztett számítási teljesítményigény exponenciálisan növekszik (Mehonic–Kenyon 2022), és a technikai megvalósítás legemlékezőbb követelményeit sem tekintő matematikai eljárások (pl. mélyrétegű tanulás) fejlesztése egyértelműen felelőssé tehető érte. A szükséges energiaigény nagyságrendje már rövid távon is a Földön megtermelhető energia összes mennyiségének nagyságrendjét közelíti (Bailey 2022). Emellett az MI által előidézett szénkibocsátás, vízfelhasználás, környezetszennyezés, disszipáció stb. növekedése jó úton van ahhoz, hogy az emberiség fennmaradásának egyik legnagyobb akadályát képezze (URL8). Lényegében erre utal, hogy *„a gépi tanulás olyan modellt használ, ami költséges, alacsony hatékonyságú és fenntarthatatlan”* (Bailey 2022).

Nagy bátorságra (vagy botorságra) vall egész iparágakat (lényegében jövőnket) MI-alapúvá alakítani, mégpedig kötelező jelleggel, szakmai megalapozottság nélkül, politikai akaratként megfogalmazva. Ami azzal együtt, hogy egyre csökkenő számítási határfokkal, egyre növekvő energiapazarlással, vízfelhasználással, környezetszennyezéssel és szénkibocsátással alkalmazzuk, elgondolkodtató; különösen, hogy az adatmennyiség racionalizálása fel sem merül.

Az MI hullámai

Az MI neve lényegében azokat az elvárásokat tükrözi, amiket első kutatói az 50-es években feltételeztek képességeinek határaitól: az emberi intelligencia gépi megvalósítása, főként gondolkodás, *magas szintű kognitív képességek modellezése volt a cél* (human-imitating) (Jordan 2019), a társ-tudományok eredményeinek felhasználásával. A vonatkozó társ-

tudományok azonban csak alacsonyabb (nem kognitív) szinten foglalkoztak jelfeldolgozással és döntéshozatallal. Rögtön az elektronikus számítógép megalkotása után megszülettek a túlzó elképzelések a *„gondolkozó, tanuló és kreatív gépek”* építésére vonatkozóan, bár megalkotói *pszichológiai modellezésre és nem mérnöki eszközökkel létrehozott berendezések készítésére szánták* (a más tudományterületen történő „túlhasználat” újabb példája). Még arra vonatkozóan is eléggé korán megfogalmazódtak kérdések (Turkle 1991), hogy az MI valóban olyan jellegű-e, mint amelyet az emberi agy is használ problémák megoldására.

A megjelenő új technológiákkal kapcsolatos várakozás a tapasztalat szerint nagyjából ugyanúgy változik; a technológiai várakozás görbéje, avagy hype néven ismert. Megjelenésük után óriási a várakozás, egyre több terület fedezi fel azokat és alkalmazza saját területére. Idővel viszont leáll az újabb területek mohó belépése, továbbá a régebben elkezdett alkalmazási területek nem érik el az elvárt (nem feltétlenül reális) sikereket, és a csúcs után következik a kiábrándulás időszaka. Ezután következik a megértés és letisztulás, hogy voltaképpen mire jó és mire nem jó, továbbá az egyes területeken mire képes az új technológia, és ezután valamilyen szinten stabilizálódik használata.

Az MI, mint technológia részben ezzel egyező, részben ilyen értelemben szokatlan viselkedésű. Több hullámban is megszületett (minden szakember generáció újra felfedezte és az éppen rendelkezésre álló technológiával nagyrészt függetlenül megalkotta), de egyes hullámai a hype általános viselkedésének megfelelnek. Az MI hullámaiban azonosíthatjuk a *„felfokozott várakozás”* csúcseit, amiket törvényszerűen követnek a csalódás hullámvölgyei. Természetesen eltérő technikai és társadalmi körülmények között, az előző hullámból származó *„konszolidált használat”* alapvonala fölött.

Az első hullám lényegében egyrészt a biológiai és elektronikai működés vélt hasonlósága és a logikai működés gépesítése igényének megvalósításából fakadt, de hamar kiderültek a megvalósítások technikai, sőt elvi korlátai. A második hullám kifulladását a társ-tudományoknak tulajdonították (nem elég fejlett elektronika, nem eléggé jó matematikai módszerek, az agy működésének nem kellő szintű ismerete stb.). A mostani hullám már teljesen más hátteret talált: minden területen jelentős fejlődés következett be. Az általános számítógépesítés miatt megnőtt a számítógépes feldolgozás elterjedtsége és szerepe. Mivel kezelhetetlenül nagygyá duzzadt a számítógépekkel termelt és általuk feldolgozandó adatmennyiség, és egyrészt nem is áll rendelkezésre megfelelő mennyiségű és szakértelmű programozó szakember, másrészt igen drága a szakértelmük, felmerült az ötlet, hogy *matematikai módszereken alapuló, megfelelően általános*

szoftver használatával, egyedi programozás nélkül dolgozzák fel az adatokat.

Az általános feldolgozási módszerek működési hatásfoka mindig lényegesen kisebb, mint a speciálisaké; az MI esetében ez akár nagyságrendeket is rontott a feldolgozás hatásfokán. A (főként matematikus) szakértők azonban nem különböztették meg az energiafogyasztással és szén lábnyommal valóban nem rendelkező MI-t, és az ezeket a matematikai módszereket alkalmazó gépi intelligenciát, így az ezzel járó problémákat (amelynek nagy részét a fejlett nyelvi modellek generálták) csak a gépi intelligencia tömeges alkalmazása hozta felszínre.

A várható tömeges felhasználást természetes nyelvet használó interfészek kifejlesztésével akarták segíteni. Ennek segítségével fejlesztették a nyelvi modelleket, amelyek használata később önállósította magát és mára a bajok fő okává vált. Továbbá az internet elterjedtsége miatt jelentősen megnőtt eme adatok elérhetősége és a társadalom (beleértve a társ-tudományokat is) várakozása az MI elvárt jövőbeli teljesítményére vonatkozóan. Ne feledjük el a kockázati tőke profit várakozását sem.

Az MI mostani hulláma kifulladásának 2023 közepétől vagyunk tanúi; az előző két nagy hype alapján, törvényszerűen. A felfokozott várakozás visszaütött. Az „intelligencia” szóhoz fűződő képzetek, párosulva a szakterületek legalábbis merész jövőképeivel, irreális elvárásokat, reményeket és félelmeket generáltak; éppúgy, mint az előző két hullám idején. Ironikus módon, éppen a használat megkönnyítésére fejlesztett nyelvi interfész tette nyilvánvalóvá, hogy az elvárások irreálisak. Mára már a napi sajtó is foglalkozik a felfokozott várakozásból eredő csalódottsággal. A kezdeményezők, támogatók, javaslattevők, gyártók, fejlesztők stb. egymással versengve igyekeznek kihátrálni az MI mögül; azért ügyelve, hogy ne veszítsék el arcukat. Egy dolog nem változott az MI körül: *nem értik valójában mi is, mire való, hogyan működik, mire jó és mire nem jó az MI.*

A gyakorlati (pénzügyi, tudományos, technológiai stb.) hasznót remélt gazdasági szereplők csalódtak. Az ilyen szempontokat nem tekintő, csupán szórakozni vágyó vagy házi feladatának megoldási kötelezettségét teljesíteni akaró stb. nagyközönséget azonban még (vagy egyre inkább?) hatása alatt tartja a varázs, és az MI alkalmazásába befektető gazdasági szereplők futnak a pénzüik után. A technikailag megvalósítható *gépi intelligencia* (bármilyen nagy mértékben tökéletesített eszközökkel megvalósítva is) *csupán nevében mutat hasonlóságot az emberi intelligenciához; ezért a kutatások nagy része megvalósíthatatlan célt tűz ki. „A gépi tanulásban óriási adatmennyiség és feldolgozó kapacitás használata sem nem hatékony, sem nem jövő-*

biztos. Az MI-nak sokkal okosabbá és nagyságrendekkel hatékonyabbá kell válni” (Webber 2022).

A nyilvános és nagymértékű kiábrándulás oka lényegében a (z LLM alapú) ChatGPT, de a fogalmi tisztázatlanság miatt az egész MI került gyanúba. Nem kellene azonban a fürdővízzel együtt a gyermeket is kiönteni. A betiltás után az EU drasztikusan (bár megértés hiányában, igen rosszul) módosította az MI-vel kapcsolatos álláspontját, ami nyilván maga után vonja Magyarország Mesterséges Intelligencia Stratégiájának (URL4) hasonló értelmű módosítását. A technológiai várakozás görbéjének szokásos menete szerint néhány hónap múlva következhet(ett volna) a konszolidált használat időszaka. Nagy kérdés, hogy a fejlesztők önként vállalt önkorlátozása, a felhasználók fokozódó óvatossága és a hatóságok addig meghozott szabályozásai után hogyan folytatódhat a történet. Annak, hogy az említettek megértenék, hogy mit fejlesztenek, használnak és szabályoznak, semmi jele.

Az MI kategóriái

Bár ma már semmi nyoma annak, hogy emlékezzenek rá, 1959 óta létezik (Samuel 1959) az MI-vel kapcsolatos módszerek matematikai alapú felosztása, amelynek használata sok bajt megelőzhetett volna. Ebben a szerző a gépi tanulási módszereket lényegében három kategóriára osztja. A *determinisztikus függvényosztály* jól meghatározott eredményt ad, mint pl. a szorzótábla vagy egy játékban következő legjobb lépés, vagy szükséges vízszint. A *probabilisztikus függvényosztály* valamely jól ismert tulajdonságú eloszlás egy elemét adja eredményül, pl. egy szabályozott mennyiség egy bizonyos érték körül, egy bizonyos szórással rendelkező eredményt ad, vagy egy idősor várható következő eredményét megalapozott matematikai módszerekkel megbecsüli. A *generatív függvényosztály* látszólag hasonló a probabilisztikushoz: ott is valamiféle eloszlás alapján adunk eredményt, de az eloszlás tulajdonságait félig-meddig a felhasználó és a felhasználási körülmények szabják meg, azaz *lényegében semmit nem tudunk róla előre mondani.*

A jelenleg problémákat okozó ChatGPT generatív modellt használ (ezt jelenti nevében a ‘G’). A matematikában kevésbé ismerősök számára, némiképp egyszerűsített magyarázatként: a három függvénnyel, három különböző módon határozzuk meg a $2+2=?$ feladat eredményét. A feladat megoldását a *determinisztikus* esetben a szokásos módon, egzakt matematika használatával modellezzük. Az eredmény többszöri próbálkozás esetén is ugyanaz, jól tesztelhető.

Probabilisztikus esetben gondoljunk pl. a liszt csomagolására. A 2 kg-os csomagolás esetében szinte biztosak lehetünk abban, hogy a mérlegre téve nem 2000 g lesz a mutatott érték a mérlegen, és biztosak lehetünk

abban, hogy sok mérés esetén átlagosan megkapjuk a 2000 g értéket, valamilyen kis szórással. *Ismerjük az eloszlást* és annak tulajdonságait, továbbá elfogadjuk, hogy az esetek néhány %-ában mondjuk csak 1997 g-os zacskót kaptunk. Mind az 1997 g-os, mind a 2002 g-os zacskó esetén az mondjuk, hogy vettünk egy 2 kg-os csomagolást lisztet. Az eredmény minden egyes próbálkozás esetében más; csak a próbálkozások eredményének sokaságáról tudunk statisztikus megállapítást tenni; csak ilyen értelemben tesztelhető.

A generatív eloszlás és annak Turing-tesztje

Már a ChatGPT-re gondolva, részletezzük a generatív eloszlás néhány tulajdonságát. A fogalom megértéséhez gondoljunk a $2+2=?$ feladat kérdőíves közvélemény-kutatással (vagy éppen egy közösségi fórumon feltett kérdésre adott válaszok összegzésével) történő megoldására.

Első lépésként megkérdezzük az embereket, hogy véleményük szerint mennyi a helyes eredmény. Mondjuk az emberek 30%-ának válaszában a helyes érték 4, 35% szerint 3, 25% szerint 5 és 10% szerint 6. Más közvélemény-kutatás és más megkérdezettek véleménye alapján, más tapasztalati eloszlást kapunk. Akárhogyan is, kapunk egy generatív (tapasztalati) eloszlást, amelyik félkész állapotban tartalmazza a feladat megoldását. *Nem ismerjük előre magát az eloszlást, de meg kell adnunk, hogyan készítünk belőle választ.* Van ugyan átlagértékünk is, meg szórás is, meg az első függvény módszere alapján ismerhetjük a helyes értéket is; de mi lesz a feladat eredménye?

Az eloszlás átlagértékét nem adhatjuk meg eredményként, mert az eredmény csak egész szám lehet. Ezért adjuk meg a kérdőíven feltett kérdésre kapott válaszok valamelyikét (vegyük észre: feltételezzük, hogy a helyes válasz is közöttük van). Első próbálkozásként vegyük a leggyakoribb választ: $2+2=3$. Az így megadott válasz ugyanezzel a módszerrel, egy másik felmérés szerint, más érték is lehet, a mi programunk szerint viszont a mi kérdőíves betanításunkkal létrehozott adatkészlet használatával a helyes válasz mindörökké 3. Azaz, programunk minden esetben hibás választ ad. Az algoritmus ugyan jó, csak rosszul választottuk meg a betanításhoz használt adatkészletet.

A *probabilisztikus* modellhez való hasonlóság alapján mondhatjuk azt is, hogy válasszuk az adatkészlet valamelyik elemét válasznak. Sem a kérdezőbiztosnak, sem a generatív modellnek, sem a válaszadónak nem kell tudnia, hogy erre a kérdésre csak egy helyes válasz van. Végül is, egy „intelligens” társadalomban többféle helyes válasz is létezhet, sőt, a társadalom valamelyik intelligens tagja is adhat a körülményektől vagy az időtől függően más válaszokat ugyanarra a kérdésre.

Azaz, valamilyen módszerrel, alkalomszerűen kijelöljük a kérdőíven megadott valamelyik választ, és azt adjuk meg válaszul a feltett kérdésre. A válasz a TI egy képviselőjétől érkezett, tehát a kérdező a TI által adott választ kap; azaz olyat, amelyet egy találmányra kiválasztott embertől is kaphatna. A program tehát pontosan ugyanúgy viselkedik, mint ahogyan találmányra kiválasztott emberek viselkednének. Teljesíti a Turing-tesztet: nem tudjuk megmondani, hogy a válasz közvetlenül egy embertől vagy a válaszokat összesítő programtól érkezett.

Nem melleleg: ha a helyes érték is a lehetséges válaszok között van, a mi programunk már nem nulla valószínűséggel ad helyes választ. Azaz, valakinek rögtön az első esetben, másnak akkor sem, ha tízszer megkérdezi. Sőt, matematikailag korrekt értelemben bejelenthetjük, hogy a mi programunk 30% valószínűséggel helyes eredményt ad, pontosan ugyanolyan arányban, mint a TI. Azaz sikerült gépi módon előállítanunk az intelligenciát, ami *egyetlen* feladatban teljes körűen, Turing értelmében imitálni tudja az emberi intelligenciát. Igaz, hogy csak egyetlen feladat megoldására, de ha hozzáadjuk a $3+2=?$ feladat megoldásának kérdőívét is, akkor már két feladatra is kiterjed. További kérdőívek hozzáadásával előbb-utóbb eljutunk oda, hogy a „feladatok széles körében” tudja imitálni az emberi intelligenciát. Lényegi különbség azonban nincs.

Ha elég buzgók vagyunk, akkor a generatív függvényt akár „kreatívva” is tehetjük. Az ugyanis akár véletlen is lehet, hogy a helyes válaszok a 3, 4, 5, 6 számok közül kerülhetnek ki (igaza van, ez az adott kérdőív kitöltőitől függ: nem felelet választós). Miért ne lehetne akár 2 vagy 7 a helyes válasz? (néhányek ezt az „intuíción” fogalommal tévesztik össze, mások egyszerűen kitalálásnak, hallucinációnak tekintik.) Azt ugyan először szomorúan vesszük tudomásul, hogy a $3+3=?$ feladat kérdőíve még továbbra sem szerepel programunk tudásában, de a kreativitás engedélyezésével már erre a kérdésre is kapunk választ. Végül is, látjuk, hogy két számról, azok összeadásáról van szó, meg két kérdőíves adatkészletről, továbbá a kreativitás is megengedett. Így az utóbbi kérdésre is rávágathatjuk azonnal a választ. Még akár az is előfordulhat, hogy helyes választ ad a programunk, legfeljebb a találati pontosság kisebb lesz. Viszont nem kell beismernünk, hogy csak bizonyos számok összeadásáról van ismeretünk (legfeljebb azt, hogy csak összeadásról). Baj csak akkor van, ha a kérdező tudja a helyes választ. Különben abban a meggyőződésben kel fel a képernyő elől, hogy programunk mindenre tudja a választ (legalábbis úgy csinál). Ha egy nem-tanítható változatról van szó, akkor ez a verzió megbukott. (Bár az előbbieket szerint ez nem biztosan derül ki: ha a készítőik szerint 30%-os pontossággal működik, akkor a

nyilvánvalóan hibás válasz ellenére is el kell fogadnunk, hogy maga az MI rendszer jó. Sőt, még az emberi tévedést is imitálni tudja).

Eddig „MI algoritmus”-unk kizárólag az eloszlás kezelésére szorítkozott. Ha azt látjuk, hogy sokan tesznek fel ilyen jellegű kérdéseket, akkor készítünk egy „MI algoritmus fejlesztést”. Kiegészítjük az algoritmust azzal, hogy készítünk egy összeadó szubrutint (a determinisztikus függvény szerint), és ha két egész szám összeadását kéri a felhasználó, akkor megértjük a két számot, elvégezzük az összeadást, megadjuk az eredményt, és az újabb verzió már átmegy a vizsgán (teljesíti a benchmarkot). Sőt, tökéletesen működik. Mármint két szám összeadását tökéletesen végzi (viszont az MI használatával ágyúval lövünk verébre). A javulás nyilvánvaló: ha többnyire ilyen feladatokat adunk, akkor jelentősen javul a rendszer pontossága. Igaz, hogy a TI két ponton is belenyúlt a generatív módon működő MI rendszer működésébe: felismerte, hogy milyen célra akarjuk használni az MI rendszert, és annak algoritmusába beépítette a TI által, a determinisztikus módszer alapján működő összeadó szubrutint. Aztán a bemutatón azt mondjuk, hogy mindezt az MI csinálja, hiszen javítottunk az MI algoritmuson.

Az adatkészlettel kapcsolatos probléma egyúttal rámutat a tanulás egyik nehézségére is. Ha folyamatosan tanuló rendszert használunk, amelyik a kapott válaszokat beépíti adatkészletébe, akkor a tudatlan vagy maliciózus felhasználók válaszai vagy akár a gépi tanulás idő- és sorrend viszonyai (lásd Függelék alapján (URL1)), az adatkészleten alapuló tapasztalati (generatív) eloszlás függvény folyamatosan változik, és egy idő után még jó eloszlásfüggvénnyel indítva is elkezdhet rossz válaszokat adni, éppen a tanulás következtében. Esetleg éppen a bemutatón siklik ki (lásd URL9). A generatív eloszlásfüggvény tulajdonságait nem ismerjük előre (ellentétben a probabilitással); azok a „kérdőíveken” megadott válaszoktól függenek, tehát esetlegesek. Időben változhatnak, manipulálhatók is.

Vegyük észre szomorú tanulságként: generatív eloszláson alapuló MI programunk:

- csak olyan eredményt tud adni, amelyet adatként már látott betanítása során
- ha mégis újat alkot, abban nincs sok köszönet
- még ha ismeri is a helyes választ, nem biztosan azt adja meg
- ha pontosabbá akarjuk tenni, TI bevonására van szükség
- még a betanítása során már látott feladatot sem biztosan tudja megoldani
- fogalma sincs arról, mit is csinál: számokkal és valószínűségekkel dolgozik.

Aki az MI generatív módszerét használja, az vállalja a fentieket és magára vessen az eredményért.

A generatív eloszlások alkalmazása médiára

Eloszlásokat generálni természetesen nem csak számokból és kérdőívek alapján lehet. Betűk, szavak, mondatpanelek; mozdulatsorok elemei; zenei hangok; esetvonások; hangulati elemek megjelenése: bármi, amit matematikailag többé-kevésbé jól le tudunk írni. (A képek/hangok manipulálása és hamisítása különösen veszélyes. Nem csak egy választási kampányban lehet el nem hangzott mondatokat valamelyik jelölt szájába adni vagy meg nem történt rendezvényen felvett képeket készíteni. Kérdéssé teszi az eddig bizonyító erejű hang és/vagy képfelvételek felhasználhatóságát is. A hangszín, a hangsúly, a hanghordozás, az akcentus, a gesztusok, a mimika stb. mind leírhatók matematikailag. Egy esetleg évek múlva sorra kerülő tárgyaláson nem csak a bíróság, hanem maga a szereplő is elbizonytalanodik: a megtévesztés a halványuló emlékekkel kombinálva tökéletes illúziót kelthet.) Az MI jelenleg legtöbb vitát kiváltó és sok veszéllyel fenyegető területe az a természetes nyelvi interfész, amelyet eredetileg a számítógépekkel való kapcsolattartás megkönnyítésére fejlesztettek, de ami azóta önálló életre kelt. Egy számítógéppel kommunikálni önmagában is nagy kihívás és az említett Turing-tesztnek is lényeges része. Részben ezek voltak az okai a természetes nyelvi interfészek fejlesztésére irányuló erőfeszítéseknek, amelyek az eredetihez képest teljesen más irányba fejlődnek. Ma már főként szöveggenerálásra irányulnak az erőfeszítések, főként arra, hogy valóságosnak hassanak (azaz minél jobban megtévezzék az MI rendszer felhasználóját), sokkal kevesebb gondot fordítva arra, hogy a tartalomnak valós értelme legyen. Bár ezt nem igazán szokták az adatvezéreltnek beállított MI rendszerek működésének ismertetésekor hangsúlyozni, a rendszerek adatait feldolgozandó, kézzel (TI által készített) algoritmusokat is be lehet és kell tenni; ezek lesznek az „MI-algoritmusok”. Hogy ezek mennyit számítanak, arra jó példa, hogy a *GPT-4* és a *GPT-3.5* még éppúgy a 2021 szeptemberi adatkészlettel dolgozik, mint a *GPT-3*; ennek ellenére lényegesen „intelligensebbnek” tűnik. Igen, az algoritmusába rengeteg új TI került. De még így is kell hozzá „prompt engineering”.

A nyelvek természetes módon tagozódnak hangokra, szavakra, kifejezésekre; vannak felismerhető szabályok, hangulati- és stíluslemek. Könnyen lehet statisztikát készíteni, hogy milyen hangok vagy szavak fordulnak elő együtt nagy valószínűséggel, és annak alapján csupán valószínűségeket használva, értelmesnek tűnő (mert gyakran így használt) pl. szókombinációkat létrehozni.

Ami önmagában igen nagy számú lehetőséget jelent (pl. az angolban a kb. 40,000 szóból előállított, csupán kétszavas kombinációkkal milliárdnyi lehetőséget kellene vizsgálnunk), amiből azonban bizonyos lehetőségek kizárásával (erre szolgálnak azok a bizonyos nyelvi modellek) észszerű nagyságú kombinációkra korlátozzuk a lehetőségeket. Általában azonban még ezekre sem tudunk egyszerű matematikai algoritmusokat megadni. Ilyenkor próbálunk meg neurális hálózatokat használni a feladat megoldására.

A neurális hálózatok abból az elképzelésből indulnak ki, hogy természetes neuronjaink hálózata matematikai algoritmus nélkül tud megoldani feladatokat, így hasonlóan gondolt mesterséges neuronhálózatokat hoznak létre, amelyekben az elemek egymással kommunikálnak, rétegeket képeznek, működési paramétereiket állítják (erről jó és érthető összefoglalót ad Wolfram (2023)). Hogy miért pont úgy, ahogyan, arra mindig akad egy olyan feladat, amelyre azt lehet mondani, hogy „mert így működik”. Másik feladatra már nem (lásd például a $2+2=?$ feladatot), így egy szűk körű intelligenciát lehet megvalósítani. Ha egy nyelvi feladatra szélesebb körű „intelligenciát” akarunk létrehozni, akkor igen nagy számú (nyelvi modellek esetében milliárdos) mintát kell használnunk a betanításhoz; a megvalósítás hosszadalmas (akár több hónapos) és a siker esetleges. A nyelvi modellek már önmagukban kizárják a lehetetlen kombinációkat, a gépi megvalósítások pedig az ún. gyorsítók használatával tovább csökkentik a megoldások előállításához szükséges időt, a megoldás pontosságát csökkentve. A kereskedelmi szolgáltatók pedig a leggyakoribb kérdésekre adandó válaszokra kész szókombinációkat vagy akár mondatpaneleket állítanak össze, amiket korábban (még a felhasználói kérdés elhangzása előtt) olyanra formálnak, amelyet a TI is használna. Az elvből következően az MI válaszainak összeállításakor semmi olyan alkotóelem nem kerül a válaszba, ami korábban nem szerepelt a betanításhoz használt szövegekben, és annak nagy része már a kérdés elhangzásakor össze volt állítva. A mondatok és kifejezések összeállítása előfordulási gyakoriság alapján történik, de az ún. MI algoritmusok is szerephez jutnak. A fejlesztés arra irányul, hogy a válasz formailag minél valóságosabb legyen; szerencsés esetekben a használati gyakoriságok valamennyire tartalmi összefüggés látszatát is kelthetik. Az egyhangúság csökkentésére változnak a szavak és a szókapcsolatok, a hiányzó tényeket pedig „gyakran együtt használt” alapon másikk (látszólag akár oda illő) ténnyel pótolják.

Javasolt osztályozás

Amint fentebb láttuk, a generatív eloszlások használatakor „orosz rulett”-et játszunk: maga a fejlesztő

sem tudja megmondani, mi lesz a feltett kérdésre a válasz. Az okát abban találtuk meg, hogy *matematikailag nem ismert tulajdonságú eloszlást használtunk, a válasz kiszámítási algoritmusát az eloszlás ismerete nélkül kell elkészítenünk, és valószínűségi elemek használatával adjuk meg az aktuális választ.* Ez a hatás tisztán elméleti felhasználásban, a csupán matematikát használó rendszerekben is jelentkezik. A technikai megvalósítás további bizonytalansági tényezőket hoz be (lásd Függelékben (URL1)), amelyek kvázi-véletlenszerű működést eredményeznek, gyakorlatilag teljesen átláthatatlanná téve az ún. MI-rendszerek működését; fejlesztők és felhasználók, tanulni és szórakozni vágyók számára egyaránt; beleértve a rendszert munka céljára használni kívánókat is. Felidézve, hogy az MI fejlesztés eredeti célja az emberi viselkedés modellezése volt, az látjuk, hogy a mostani fejlesztés teljesen félreszaladt: inkább irányítani akarja az emberi viselkedést, mintsem hasonlítani rá. A káosz akkora, hogy a kognitív tudomány (ami az eredeti célközönség volt) az AI rövidítést az „Alien Intelligence” (földön kívüli intelligencia) jelentéssel használja (Frank 2023). Már szó nincs arról, hogy emberi intelligenciát más technikai megoldással hozzunk létre.

A vágyalmok alapján lényegében három kategóriára lehet osztani a megvalósításokat. A legtágabbról, az emberi intelligencia teljes körét megvalósító (sőt, a fanatikusok szerint: felülmúló) technikai rendszerekről legfeljebb a realitáshoz egyáltalán nem kötődő rajongók beszélnek. A legszűkebb, az emberi intelligencia megnyilvánulásának egyetlen, vagy csak nagyon kevés vonatkozását megvalósító rendszerek (rutinfeladatok folyamatos végzése, szoros szabálykövetés, lehetséges esetek felmérése, osztálybasorolás, nagy mennyiségű számítás gyors elvégzése) akár az embernél jóval gyorsabbak lehetnek, hibátlanul dolgoznak, nem fáradnak. A középső, az előbbinél jóval ambiciózusabb rendszerek többre törekszenek, de képességeik és eredményeik egyáltalán nem állnak arányban kitűzött céljaikkal (a jelen történések leginkább az ebbe a kategóriába tartozó rendszerekre vonatkoznak). Nem véletlenül, ugyanaz történik, mint amit az MI első hullámát lezáró jelentés (URL7) említ: a kevéssé sikeres és perspektivikus területek igyekeznek a sikeres és perspektivikus területek álcája mögé bújni, és az „egységes MI” kép alkalmazásáért harcolni. Ami törekvés a „vezetői összefoglaló” szintjén sikeres is. Ezeknek a területeknek érdeke, hogy a működés „fekete doboz” (URL10) jellegű legyen: ha tudjuk, mi és hogyan történik, el tudjuk dönteni, hogy az jó-e nekünk. A közvetlen üzleti érdekek indokolják, hogy a fejlesztett feldolgozási algoritmusok, a válaszadáshoz használt adatkészletek és a felhasználó által kapott válaszok előállításának módja mind céges titoknak számítanak.

A felhasználási területek alapján lehet talán a legrosszabbul és legkevésbé megalapozottan osztályokra bontani az MI-t, és csupán a tiltás fokára vonatkozó kategóriákba sorolni az alkalmazásokat, mint ahogyan az EU jelenleg tette (URL11). Egy jellemző példa, hogy a „magas kockázatú” csoportba sorolja a „veszélyes viselkedést ösztönző hangtámogatást használó játékokat”, de a „minimális kockázatú” csoportba az „MI-kompatibilis videojátékokat”. Közben van még két kategória. Emberek sorsára vonatkozó döntést hozni MI alapján a kormányoknak tilos, pl. magáncégeknek szabad. Teljesen kivonja a szabályozás alól a katonai, belügyi stb. alkalmazásokat. Valószínűleg amúgy sem lenne rá hatásköre. Az a tény, hogy az oktatás és ismeretterjesztés – az EU bizottsági javaslat ellenére – nem tartozik abba a kategóriába, amelyekben az MI eszközeivel valótlanságok, téves következtetések, tévinformációk terjesztése stb. tiltott, komoly kockázatot jelent; önmagáért beszél. Sokkal értelmesebb lenne a működés elve alapján részekre bontani az MI területeit, mint a felhasználási terület szerint, a működési módtól függetlenül (lásd Függelékben (URL1)).

Használhatóság

Az MI használhatóságát sok mítosz övezi. Mint említettük, lényegében azt az illúziót igyekszik kelteni, hogy emberi partner; a fejlesztés nem működésének javítására, hanem az illúzióeltérítés tökéletesítésére irányul. Emberi tulajdonságokkal úgy közelíthető, hogy egyrészt igen nagyképű (mindenhez hozzászól és részletesen kifejti véleményét, igyekszik szakértőnek látszani); gátlás nélkül kitalál nem ismert (csupán nyelvtani értelemmel rendelkező) tényeket, és azokat a valódi tényekkel összeötvözi (bár tagadja, hogy ezt tenné); nem tudja megadni a forrásokat, hogy honnét származik az információ; ismételt és konkrétabb kérdések esetén elbizonytalanodik. Valójában úgy állítja össze „mondanivalóját”, ahogyan azt feltételezése szerint a kérdező hallani szeretné. Aki valóban innét szeretne tájékozódni, azt akár teljesen félrevezetheti; ráadásul a változatosság érdekében a látszólag változatlan körülmények között (szándékoltan véletlenszerű, szándéktalanul technikailag kvázi-véletlenszerű vagy a használat idejétől és körülményeitől függő mértékben) akár homlokegyenest eltérő válaszokat adhat. Ha az érdeklődő saját maga által jól ismert tényekre vonatkozó kérdésekkel teszteli az MI tudását, két-három kérdéssel kiderítheti, hogy az imponáló válaszkészség mögött előre elkészített sablon nyelvtani elemek állnak, és a válaszoló nem ismeri az általa használt szavak és mondatok jelentését, csupán a gyakori használat alapján tippel.

A működés mögött egy korlátozott (bár elképesztően nagy) méretű adatkészlet és annak készítőinek emberi intelligenciája áll, az MI algoritmus pedig olyan, amelyet a készítés időpontjában a készítő a kérdés feltevésének körülményeire előre láttak. Amint valami hiba kiderül, a fejlesztők (TI) javítanak az algoritmuson és a javított algoritmus már teljesíti a célul kitűzött tesztet. Úgy tűnik, a skalázhatóság megértése sem erőssége az MI hívőknek. Természetesen ma mindenki, mindenre az MI-ben találja meg a csodaszert, és „minél nagyobb annál jobb” elven épült rendszereket akar használni. A fejlesztése (valószínűleg nem véletlenül) a szoftver fejlesztésnek ahhoz a módjához hasonlít, amikor először kitűzzük a teljesítendő feladatot (milyen benchmark teljesítése), aztán addig gyötörjük az algoritmust és az adatokat, amíg a tesztprogram sikeresen le nem fut. Ami természetesen csak annyit bizonyít, hogy az illető tesztet sikerrel vette a program. Egy másik teszt természetesen másik hibák kijavítását teheti szükségessé. Ha az nem sikerül, csak az első benchmark sikeres teljesítését tesszük közzé. Természetesen folyik a rakéta-ellenrakéta fejlesztés: a ChatGPT_{3.5} által még nem teljesíthető teszteket a ChatGPT₄ már teljesíti, hiszen ismert a hibák oka. Közben azonban megszületik a másként tesztelő benchmark, amely már a ChatGPT₄ számára is nehéznek bizonyul (URL12). Figyelni kell az értékelések megfogalmazására és az értékelők mögött álló (többnyire üzleti) érdekekre. Orvosi vizsgakérdések megválaszolásában a *GPT-4 lényegesen jobban vizsgázott, mint a GPT-3.5*, de még mindig alatta marad egy átlagos emberi orvostanhallgató teljesítményének (URL12). Ami nem áll ellentétben a GPT-4 technikai bejelentésével (URL13), miszerint felülmúlja az emberek többségének vizsgaképességét. Az emberek többsége valóban rosszul teljesítene az orvostanhallgatók tudásának mérésére szolgáló teszten. Bemutatunk azonban két olyan esetet, ahol a követelményszint magasabb és a benchmarkot nem lehet manipulálni. Az elsőben a természet fejleszti a benchmarkot, a másodikban a technológia.

A COVID-19 pandémia kezelése

Sajnos a közelmúltban előfordult az az esemény, ami „élesben” tesztelte az MI képességeit. Ilyen szempontból nagyon jó időzítéssel, 2019-ben, a hype (technológiai várakozás) erősen felfutó periódusában, a matematikai módszerek és az elektronikai eszközök teljességének birtokában, a módszerek és eszközök használatában jártas nagy számú kutató jelenlétében, a COVID-19 pandémia előidézte azt a társadalmi szükségletet, hogy az MI is harcba szálljon a járvány megfelelő kezelésének érdekében. Ehhez a kormányoktól megkapták a szükséges anyagi- és emberi erőforrásokat is.

Illusztratív példaként szolgál ama számos eszköz esete, amelyeket a COVID-19 járvány diagnosztizálására és előrejelzésére fejlesztettek ki, és amelyek megvalósítására 2019. és 2020. során jelentős erőforrásokat biztosítottak. Ezek analízise világos következtetéseket hozott. A nagy tekintélyű The Alan Turing Institute jelentése (von Borzyskowski et al. 2020) az UK Data Science and Artificial Intelligence közösségnek a COVID-19 járvány elleni küzdelemre adott válaszát ismerteti. Következtetése, hogy bár számos projektben széles skálán történtek fejlesztések, az eredmények nem igazán voltak használhatók, jócskán teret hagyva a jövőbeli járványok sokkal jobb előrejelzésének. Ugyanerre jutott a British Medical Journal-ban megjelent tanulmány, amely a COVID-2019 diagnózisára és prognózis céljára kifejlesztett előrejelző eszközök mindegyikét alkalmatlannak találta klinikai használatra. Mindössze két olyat talált, amelyik további vizsgálatra érdemes. Egy, a Nature Machine Intelligence folyóiratban megjelent tanulmány 415 olyan gépi tanulás alapú eszközt vizsgált, amelyek mellkasröntgen és CT felvételek alapján próbáltak a páciensek esélyeire következtetni. A következtetés megint csak nem biztató: egyik sem találtatott alkalmasnak klinikai használatra.

Úgy tűnik, ez az éles teszt, amiben a résztvevők minden lehetőséget megkaptak (a nagy figyelmet és a kiugrás/bizonyítás lehetőségét is), továbbá szó szerint az életük volt a tét, egyáltalán nem igazolta az EU 2018-as reményeit, hogy „a mesterséges intelligencia segít a világ legnagyobb kihívásainak megoldásában: a krónikus betegségek kezelésében”. Hasonló eredményre jutott a National Institute of Health is: „Valójában, még ha számos alkalmazás célozta is meg a COVID-19 előrejelzését és diagnosztizálását, közülük csak néhány elég érett arra, hogy klinikai körülmények között alkalmazni lehessen” (Comito–Pizzuti 2022).

Elektronikus tervezés

Érdekes és stílusos ötlet hibás elvű LLM-et használni a hibás elvű hardver gyorsítóegységek tervezésére (Hemsoth Pricket 2023). Az MI/LMM fejlesztésében érdekelt cégek által támogatott technikai újságírók véleménye szerint is, *még nem sokat látni abból, amit az ilyen elven működő elektronikus fejlesztő rendszerektől elvárnak.* Az eddigi tapasztalatok szerint nem boldogulnak az összetett feladatokkal, nem tudják elérni a zárt forrású (TI által készített) rendszerek által elért teljesítőképességet. A gyorsító rendszerekben használt megoldások ugyanis eltávolítják a szükséges információ egy részét, a látszólagos gyorsabb működés érdekében félbevágják a számítást (lásd Függelékben (URL1)). Ennek következtében a tervezőrendszer éppúgy

bugyután viselkedik, mint a csevegő robotok: az általuk tervezett rendszerek is bugyután viselkednek, vagy egyáltalán nem működnek. Továbbá ezek a problémák csak a TI által készített „demonstrációs célú” darabok elkészítésével csökkenthetők, ami másképpen fogalmazva azt jelenti, hogy ha betanításhoz használt adatkészletben nem szerepel a tervezendő „Intellectual Property” (IP), akkor az MI nem fogja azt megtervezni. *Újat nem tud alkotni, csak egy (vagy több) régit megtalálni és abból kicsit átfogalmazva, egy „generált” IP-t készíteni.* Hogy az működőképes-e, az attól függ. Az MI nem tudja, mi a terv célja. Csupán készít valamit „tartalmilag és formailag”, probablisztikus értelemben, ahhoz hasonlóan.

Összefoglalás

A mesterséges intelligencia néven ismertté vált számítógépes technológia egyre inkább bevonul életünkbe. Működése és hatásai egyre átláthatatlanabbá válnak, miközben fogalma is egyre tisztázatlanabb. Olyan számítógép-alapú technológia, amelynek főbb tulajdonságait megérteni csak akkor lehet, ha technikai működésének elveit legalább madártávlatból ismerjük. A tanulmány ebben próbál segíteni. Nem csodaszer, de megfelelő óvatossággal akár segítségünkre is lehet, bár manapság a fejlesztések nem erre irányulnak. Sem a hívők sokszor tapasztalt (a marketinges „vezetői összefoglaló” anyagokon alapuló) csodavárása, sem a világvége hangulat nem reális. *„Az LLM technológia használata kikerülhetetlen, ezért annak tiltása nem megoldás”; A ChatGPT és más LLM rendszerek meggyőző, de sokszor hibás szöveget állítanak elő; használatuk torzíthatja a tényeket és tévinformációkat terjeszthet” (van Dis et al. 2023). Meg kell tanulnunk együtt élni vele: mértékkel, utólagos emberi ellenőrzés mellett, az emberi tevékenységek rutin jellegű részének gyors elvégzésére használni, de semmiképpen nem helyette.*

Irodalomjegyzék

- Augusto, E.–Gambino, F. (2019) Can NMDA Spikes Dictate Computations of Local Networks and Behavior? *Frontiers in Molecular Neuroscience*, Vol. 12/238. doi:10.3389/fnmol.2019.00238 <https://www.frontiersin.org/articles/10.3389/fnmol.2019.00238/full> [Letöltve: 2024.02.16.].
- Bailey, B. (2022) AI Power Consumption Exploding. *Semiconductor Engineering*, August 15th, 2022. <https://semiengineering.com/ai-power-consumption-exploding/> [Letöltve: 2024.02.16.].
- Biever, C. (2023) ChatGPT broke the Turing test – the race is on for new ways to assess AI. *Nature*, 25 July 2023.

- <https://www.nature.com/articles/d41586-023-02361-7.pdf> [Letöltve: 2024.02.16.].
- Buzsáki, Gy. (2019) *The Brain from Inside Out*. Oxford University Press. ISBN: 978-0-19-090538-5.
- Chu, D.–Prokopenko, M.–Ray, J. C. J. (2018) Computation by natural systems. *Interface Focus*, 19, October 2018. doi: 10.1098/rsfs.2018.0058 <https://royalsocietypublishing.org/doi/10.1098/rsfs.2018.0058> [Letöltve: 2024.02.16.].
- Comito, C.–Pizzuti, C. (2022) Artificial intelligence for forecasting and diagnosing COVID-19 pandemic: A focused review. *Artificial Intelligence in Medicine*, Vol. 128:102286. doi: 10.1016/j.artmed.2022.102286 <https://www.sciencedirect.com/science/article/pii/S0933365722000513?via%3Dihub> [Letöltve: 2024.02.16.].
- Conklin, A. A.–Kumar, S. (2023) Solving the big computing problems in the twenty-first century. *Nature Electronics*, Vol. 6. pp. 464-466. <https://www.nature.com/articles/s41928-023-00985-1.epdf> [Letöltve: 2024.02.16.].
- Feynman, R. P. (2018) *Feynman Lectures on Computation*. CRC Press. ISBN 9780429500442.
- Frank, M.C. (2023) Baby steps in evaluating the capacities of large language models. *Nature Reviews Psychology*, 2. pp. 451-452. <https://doi.org/10.1038/s44159-023-00211-x> [Letöltve: 2024.02.16.].
- Hassabis, D.–Kumaran, D.–Summerfield, C.–Botvinick, M. (2017) Neuroscience-Inspired Artificial Intelligence. *Neuron*, 95(2). pp. 245-258. doi: 10.1016/j.neuron.2017.06.011 <https://pubmed.ncbi.nlm.nih.gov/28728020/> [Letöltve: 2024.02.16.].
- Hemsoth Prickett, N. (2023) What happens when LLMS design AI accelerators? *TheNextPlatform*. <https://www.nextplatform.com/2023/09/25/what-happens-when-llms-design-ai-accelerators/> [Letöltve: 2024.02.16.].
- Jones, N. (2018) How to stop data centres from gobbling up the world's electricity. *Nature*, 13 September 2021. <https://www.nature.com/articles/d41586-018-06610-y> [Letöltve: 2024.02.16.].
- Jordan, M. I. (2019) Artificial Intelligence – The Revolution Hasn't Happened Yet. *Harvard Data Science Review*, Issue 1.1, Summer 2019. doi:10.1162/99608f92.f06c6e61 <https://hdsr.mitpress.mit.edu/pub/wot7mkci/release/10> [Letöltve: 2024.02.16.].
- Kar, K.–Kornblith, S.–Fedorenko, E. (2022) Interpretability of artificial neural network models in artificial intelligence versus neuroscience. *Nature Machine Intelligence*, 4. pp. 1065-1067. doi:10.1038/s42256-022-00592-3 <https://arxiv.org/abs/2206.03951> [Letöltve: 2024.02.16.].
- Kovács, Z. (szerk.) (2023) *A mesterséges intelligencia és egyéb felforgató technológiák hatásainak átfogó vizsgálata*. Katonai Nemzetbiztonsági Szolgálat, 2023. https://www.knbsz.gov.hu/hu/letoltes/kiadvanyok/or_MI.pdf [Letöltve: 2024.02.16.].
- Macpherson, T.–Churchland, A.–Sejnowski, T.–DiCarlo, J.–Kamitani, Y.–Takahashi, H.–Hikida, T. (2021) Natural and Artificial Intelligence: A brief introduction to the interplay between AI and neuroscience research. *Neural Network*, Vol. 144. pp. 603-613. <https://doi.org/10.1016/j.neunet.2021.09.018> [Letöltve: 2024.02.16.].
- Mehonic, A.–Kenyon, A. J. (2022) Brain-inspired computing needs a master plan. *Nature*, Vol. 604/1. pp. 255-260. doi:10.1038/s41586-021-04362-w <https://www.nature.com/articles/s41586-021-04362-w> [Letöltve: 2024.02.16.].
- Pretz, K. (2021) Stop Calling Everything AI, Machine-Learning Pioneer Says. *IEEE Spectrum*. <https://spectrum.ieee.org/the-institute/ieee-member-news/stop-calling-everything-ai-machinelearning-pioneer-says> [Letöltve: 2024.02.16.].
- Samuel, A. R. (1959) Some studies in machine learning using the game of checkers. *IBM Journal Res. Dev.* 44(1959), pp. 1210-1229.
- Turkle, S. (1991) Dangerous Thoughts . . . And Machines With Big Ideas. *The New York Times*, March 17, 1991, Section 7, Page 1. <https://www.nytimes.com/1991/03/17/books/dangerous-thoughts-and-machines-with-big-ideas.html> [Letöltve: 2024.02.16.].
- URL1: Függelék <https://github.com/jvegh/TimeAwareComputing.git> [hub.io/blob/main/ChatGPT.pdf](https://github.com/jvegh/TimeAwareComputing.git) [Letöltve: 2024.03.02.].
- URL2: (2000) *Mi a MI? Avagy a mesterséges intelligencia témaköre, célkitűzései*. <https://doksi.net/hu/get.php?lid=484> [Letöltve: 2024.03.02.].
- URL3: Fang, J.–Su, H.–Xiao, Y. (2018) *Will Artificial Intelligence Surpass Human Intelligence?* <http://dx.doi.org/10.2139/ssrn.3173876> [Letöltve: 2024.02.16.].
- URL4: (2019) Magyarország Mesterséges Intelligencia Stratégiája. <https://digitalisjoletprogram.hu/files/2f/32/2f32f239878a4559b6541e46277d6e88.pdf> [Letöltve: 2024.02.16.].
- URL5: General-purpose artificial intelligence. <https://epthinktank.eu/2023/03/31/general-purpose-artificial-intelligence/> [Letöltve: 2024.02.16.].
- URL6: Jelentés a mesterséges intelligenciára vonatkozó harmonizált szabályok (a mesterséges intelligenciáról szóló jogszabály) megállapításáról és egyes uniós jogalkotási aktusok módosításáról szóló

- európai parlamenti és tanácsi rendeletre irányuló javaslatról.
https://www.europarl.europa.eu/doceo/document/A-9-2023-0188_HU.html [Letöltve: 2024.02.16.].
- URL7: Lighthill, J. (1972) *Artificial Intelligence: A General Survey*.
https://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/poor.htm [Letöltve: 2024.02.16.].
- URL8: Luccioni, A. S.–Viguier, S.–Ligozat A-N (2022) *Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model*.
<https://arxiv.org/abs/2211.02001> [Letöltve: 2024.02.16.].
- URL9: Germšek, M. (2023) How Overhyped is ChatGPT in 2023? The Truth about its Hype Explained.
<https://shape-labs.com/articles/how-overhyped-is-chatGPT> [Letöltve: 2024.02.16.].
- URL10: (2023) ChatGPT is a black box: how AI research can break it open. *Nature*, 619, pp. 671–672.
<https://www.nature.com/articles/d41586-023-02366-2> [Letöltve: 2024.02.16.].
- URL11: Európai Bizottság (2023) A mesterséges intelligenciáról szóló törvény.
<https://digital-strategy.ec.europa.eu/hu/policies/regulatory-framework-ai> [Letöltve: 2024.02.16.].
- URL12: Rosol, M.–Gašior, J. S.–Laba J.–Korzeniewski, K.–Mlynczak, M. (2023) *Evaluation of the performance of GPT-3.5 and GPT-4 on the Medical Final Examination*.
<https://www.medrxiv.org/content/10.1101/2023.06.04.23290939v1.full.pdf> [Letöltve: 2024.02.16.].
- URL13: Achiam, J. et al (2023) *GPT-4 Technical Report*.
<https://arxiv.org/abs/2303.08774> [Letöltve: 2024.02.16.].
- URL14: Európai Bizottság (2018) A közös európai adattér kialakítása felé.
<https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:52018DC0237> [Letöltve: 2024.02.16.].
- van Dis, E. A. M.–Bollen, J.–Zuidema, W.–van Rooij, R.–Bockting, C. L. (2023) ChatGPT: five priorities for research. Conversational AI is a game-changer for science. Here's how to respond. *Nature*, Vol. 614. pp. 224–226.
<https://www.nature.com/articles/d41586-023-00288-7> [Letöltve: 2024.02.16.].
- Végh, J.–Berki, Á. J. (2021) Why learning and machine learning are different. *Advances in Artificial Intelligence and Machine Learning*, 1(2). pp. 136–154. doi:10.54364/AAIML.2021.1109
<https://www.oajaiml.com/uploads/archivepdf/77931109.pdf> [Letöltve: 2024.02.16.].
- Végh, J. (2021) Revising the Classic Computing Paradigm and Its Technological Implementations. *Informatics*, 8/4(2021). 71. doi:10.3390/informatics8040071
<https://www.mdpi.com/2227-9709/8/4/71> [Letöltve: 2024.02.16.].
- von Borzyskowski, I.–Mazumder, A.–Mateen, B.–Wooldridge, M. (eds.) (2020) *Data Science and AI in the age of COVID-19*. Reflections on the response of the UK's data science and AI community to the COVID-19 pandemic. The Alan Turing Institute.
https://www.turing.ac.uk/sites/default/files/2021-06/data-science-and-ai-in-the-age-of-covid-full-report_2.pdf [Letöltve: 2024.02.16.].
- Webber, F. (2022) Third AI Winter ahead? Why OpenAI, Google & Co are heading towards a dead-end. *Cortical.io*. July 18, 2022.
<https://www.cortical.io/blog/third-ai-winter-ahead-why-openai-google-co-are-heading-towards-a-dead-end/> [Letöltve: 2024.02.16.].
- Wolfram, S. (2023) *What Is ChatGPT Doing ... and Why Does It Work?* Wolfram Media, Inc. ISBN-13: 978-1-57955-081-3 (paperback), ISBN-13: 978-1-57955-082-0 (eBook).
<https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/> [Letöltve: 2024.02.16.].

Komplex rendvédelmi MI politikai stratégiai és technológiai javaslatok a ChatGPT és más LLM-ek szempontjából

Complex law enforcement AI political, strategic, and technological proposals from the perspective of ChatGPT and other LLMs

DOI: [HTTPS:// DOI.ORG/10.53793/RV.2024.1.3](https://doi.org/10.53793/RV.2024.1.3)

Absztrakt

A rendőri szervek folyamatosan új technológiákat és módszereket fejlesztenek a bűnelkövetők felderítésére és elfogására, de az elkövetők is hasonlóképpen fejlesztik saját trükkjeiket és eszközeiket a hatóságok kijátszására. Az állam jogalkotó készsége reaktív jellegénél fogva ugyanakkor a rendvédelem általában utólag reagál az egyes jogsértésekre. A bűnözők már több ízben vették hasznát a ChatGPT nyelvi modellnek, hogy abból profitáljanak. A mesterséges intelligencia és a rendvédelem kapcsolatában számos izgalmas fejlemény várható a jövőben. Ezek között szerepelnek azok a lehetőségek és kihívások, amelyek a mesterséges intelligencia alkalmazásából és fejlődéséből fakadnak. A mesterséges intelligencia és a nagy nyelvi modellek (LLM), mint a ChatGPT rendvédelmi alkalmazása azonban számos kérdést vet fel.

KULCSSZAVAK: RENDVÉDELEM, MESTERSÉGES INTELLIGENCIA, CHATGPT, STRATÉGIA,

Abstract

Law enforcement agencies continuously develop new technologies and methods to detect and apprehend criminals, but offenders similarly devise their own tricks and tools to evade authorities. Due to the reactive nature of the legislative capabilities of the state, law enforcement typically responds retrospectively to individual violations. Criminals have repeatedly taken advantage of the ChatGPT language model to profit from it. Exciting developments are anticipated in the future concerning the relationship between artificial intelligence and law enforcement. These include the opportunities and challenges arising from the application and evolution of artificial intelligence. However, the application of artificial intelligence and large language models (LLM), such as ChatGPT, in law enforcement raises numerous questions.

KEYWORDS: LAW ENFORCEMENT, ARTIFICIAL INTELLIGENCE, CHATGPT, STRATEGY

Bevezetés

"Az embernek mindig törekednie kell a lehetetlenre, mert amit ma lehetetlennek tart, az holnap már a valóság része lehet." Ezt az idézetet általában Jules Verne nevéhez szokták társítani, ám a pontos forrás nem ismert. Verne műveiben, levelezéseiben vagy más hiteles dokumentumban nem található. Az idézet erősen hasonlít a Verne művekben kifejezett gondolatokhoz, amelyek a lehetetlennek tűnő utazások és felfedezések iránti vágyról szólnak. Nem véletlenül választottam ezt az idézetet. Ez egy csodálatos gondolat, amely a kitartás, a vízió és az emberi teljesítmény határtalanságát hangsúlyozza. Sok nagyszerű felfedezés és innováció valósult meg azáltal, hogy az emberek merészen

törekedtek olyan célok elérésére, amelyeket mások talán lehetetlennek tartottak. Az emberi képesség és technológiai fejlődés során olyan lehetőségek nyílnak meg, amelyek egykor elképzelhetetlennek tűnhettek (Iványi 2023: 58). Ez az idézet arra emlékeztet minket, hogy soha ne korlátozzuk magunkat azzal a kérdéssel, hogy mi lehetséges vagy lehetetlen. Ahelyett, hogy megállnánk az akadályok előtt, próbáljuk meg átlépni őket és tartsuk szem előtt, hogy a jövőben a mai lehetetlennek tűnő dolgok válhatnak a valóság részévé.

A nagysikerű 1984-es *Terminátor* című film folytatására több évig kellett várni, mígnem 1991-ben testet öltött egy mesterséges intelligenciával felvértezett

gyilkológép, amely immáron az embert szolgálta. Íme egyik leghíresebb ember – gép közötti párbeszéd a filmiparból:

John Connor: "Mi vagy te?"

T-800: "Én egy T-800-as típusú küldött vagyok, tervező: Cyberdyne Systems."

John Connor: "Tehát egy gép vagy?"

T-800: "Igen, a gépek közül vagyok, de nem vagyok ellened ellenségesen hangolva."

Ez a párbeszéd hangsúlyozza a film központi témáját, amely az ember és gép közötti kapcsolatot, valamint az emberiség kérdése körül forog. A film arra összpontosít, hogy vajon a gépek, különösen az önállóan gondolkodó és tanuló gépek, képesek lehetnek-e az empátiára és az emberi értékek megértésére. John Connor kérdése, ti. "Mi vagy te?" és a T-800 válasza, miszerint "nem vagyok ellened ellenségesen hangolva" azt mutatja be, hogy a gépi lényeknek is lehet emberi vonásuk és empátiájuk. A film továbbá azt is feszegeti, hogy vajon az emberi gyarlóság vagy a gépek által hozott döntések veszélyesebbek-e az emberiségre, illetve, hogy az emberiségnek hogyan kell megküzdenie a gépekkel, amelyek lehetnek segítőkészek vagy veszélyesek. A *bűnözők egy lépéssel a rendőrök előtt járnak* – szokás mondani akkor, amikor az elkövetők kiszámítják, vagy elkerülik a rendőri jelenlétet, az ellenőrzést, a nyomozást, az intézkedéseket, ennél fogva nehezebb őket kézre keríteni. Ez az idézet hangsúlyozza azt a folyamatos versenyt, amely az elkövetők és a rendőrség között zajlik. A rendőri szervek folyamatosan fejlesztik és alkalmazzák az új technológiákat és módszereket a bűnözők felderítésére és elfogására, de a bűnözők is hasonlóképpen fejlesztik saját trükkjeiket és módszereiket a hatóságok kijátszására. Az idézet azt is hangsúlyozza, hogy a rendőri szervezeteknek és a törvényen kívülieknek egyaránt folyamatosan fejlődniük kell, hogy lépést tarthassanak a másik féllel, valamint azt, hogy a bűnüldözés soha nem áll meg. Ez a folyamatos verseny és kreatív gondolkodás mindkét oldalon, az igazságügyi rendszer és a bűnözők között hozzájárul ahhoz, hogy a rendőri munka izgalmas és kihívásokkal teli maradjon. Az állam jogalkotó készsége reaktív jellegénél fogva ugyanakkor a rendvédelem általában utólag reagál az egyes jogsértésekre. A bűnözők ugyanis már több ízben

vették hasznát a ChatGPT nyelvi modellnek,¹ hogy abból profitáljanak. Ezzel szemben a rendvédelmi szervek sem ülnek tétlen. 2023. március végén az Europol innovációs laboratóriuma egy új Tech Watch gyorsjelentést tett közzé, amely a nagy nyelvi modelleket, mint például a ChatGPT-t és azok bűnüldözésre gyakorolt hatását vizsgálja. A jelentés olyan szakértői műhelytalálkozók eredményein alapul, amelyek során egy nagy nyelvi modellt, nevezetesen a ChatGPT viselkedését tesztelték, amikor potenciálisan rosszindulatú használatot észleltek (URL₂). A mesterséges intelligencia az egyik meghatározás szerint a gépek emberhez hasonló kognitív képességeit jelenti, mint például a tanulás, a tervezés, az érvelés és a kreativitás. Ezek a képességek lehetővé teszik a gép számára, hogy érzékelje környezetét, problémákat oldjon meg és konkrét cél elérése érdekében tervezze meg lépéseit. A számítógép nemcsak adatokat fogad, hanem fel is dolgozza azokat és reagál rájuk (Kiss 2023: 212). Más megfogalmazásban: a mesterséges intelligenciát tekinthetjük az egyik legújabb tudományterületnek, amely mindenki számára mást jelent. Az intelligens eszközök, az érzékelők és az emberek közötti folyamatos interakció eredményeként egyre több adatot állítunk elő, dolgozunk fel és tárolunk (Mazsu 2019). A mesterséges intelligencia (a továbbiakban: MI) és a rendvédelem kapcsolatában számos izgalmas fejlemény várható a jövőben. Ezek között szerepelnek azok a lehetőségek és kihívások, amelyek az MI alkalmazásából és fejlődéséből fakadnak. Ugyanakkor az MI rendvédelmi alkalmazása számos kérdést vet fel. Tanulmányomban az MI – kiemelten a ChatGPT – rendőrségi alkalmazhatóságát vizsgálom.

Rendészeti stratégiák

A jól megtervezett stratégia segít a rendvédelmi szervezeteknek abban, hogy hatékonyabban reagáljanak a kihívásokra és javítsák a lakosság biztonságérzetét. „Az ENSZ jogállamiság-indikátora az intézményi megközelítést alkalmazva a rendőrség, az igazságügyi szolgáltatás és a börtönök állapotát vizsgálja. Adatforrásai: adminisztratív adatok; szakértők felmérései; közvélemény-kutatások; dokumentumok felülvizsgálata. A három intézmény foglalja magában a 135 indikátort, nevezetesen a rendőrség (41 indikátor), az

kiderült, hogy a vállalat idén május óta illegálisan szerzett számos online videófiókot. A csoport a ChatGPT-n keresztül hamis videókat készített mielőtt közzétette volna a közösségi média platformjain, hogy forgalmat szerezzen és bevételt generáljon. A mai napig a banda illegálisan vásárolt több mint 1.500 videófiókot, és több mint 3.000 hamis videót adott ki. A három gyanúsítottat a rendőrség köznyugalom elleni bűncselekmények elkövetése miatt tartja fogva, az ügyet tovább vizsgálják.” (URL₁)

¹ „2023 június 02-án egy online ellenőrzés során a Shangyu kerületi rendőrség azonosított egy "Xianyijieshuo" néven emlegetett netezőt, aki videót tett közzé a Shangyu Ipari Parkban keletkezett tűzről. Rövid idő alatt a videó megtekintéseinek száma gyorsan nőtt. A hatóságok a helyszínre siettek ellenőrizni a helyszínt, ahol kiderült, hogy a videó hamis, a tűz nem volt valós. A Shangyu rendőrség gyorsan azonosított egy technológiai csoportot egy másik tartományban, amely félrevezette a hatóságokat. Június 5-én a rendőrség három gyanúsítottat fogott el. A nyomozás során

igazságszolgáltatás (51 indikátor) és a börtönök (43 indikátor)” (Kiss 2016: 7). Eszerint nemzetközileg kialakult eljárásrend alapján szorosan összefügg a rendőrség működése az adott társadalom politikája irányába tanúsított elégedettséggel. Amennyiben a rendőrség kevésbé hatékony, az negatívan hat az adott társadalom tagjaiban a szervezet iránt érzett bizalom tárgyában.

A rendészeti stratégiaalkotásról

A rendészeti stratégia megalkotásának első lépéseként szükséges megvizsgálni a szakirodalom stratégiára irányuló fogalmkörét. A stratégia fogalmi meghatározására számos definíció született. Ezek közül a legegyszerűbb: egy szervezet hosszabb távú terve bizonyos célok elérése érdekében. Fogalmazhatunk úgy is, hogy a stratégia nem más, mint egy krízishelyzetben elkészített dokumentum, amelyben a szervezet átalakításának, további vezetésének, fejlesztésének legfontosabb, meghatározó tennivalói vannak. A rendészettudománynak otthont adó katedra vezetéselméleti definíciója szerint viszont a stratégia „a szervezet jövőbeni céljaira és azok megvalósítására vonatkozó elképzelések összessége, egyben egy meghatározott vezetési folyamat. A stratégiai vezetési folyamat elemzési, döntési és megvalósítási szakaszokból áll (tervezés és végrehajtás). A stratégiai tervezés feladata a célok kitűzése és az azok eléréséhez vezető út megjelölése egy előzetes helyzetértékelés alapján.

A terveket időtartamukat tekintve több csoportra bontjuk:

- hosszú távú tervek: öt évet meghaladnak, esetleg tíz éven is túlnyúlnak,
- középtávú tervek: kettő-öt évre szólnak,
- rövid távú tervek: egy-két hónaptól egy évig terjednek.

Jellegüket tekintve a tervek lehetnek:

- stratégiai tervek: a szervezet egészére vonatkoznak, azok lényegét érintik, és működését döntő módon befolyásolják;
- taktikai tervek: egy adott részterület szervezeti egysége elérni kívánt eredményét rögzítik;
- operatív tervek: egy konkrét esemény megoldására irányulnak” (Fórizs 2014: 115).

A rendészeti stratégiaalkotás egy összetett folyamat, amely során számos különböző tényezőt kell figyelembe venni. A stratégiaalkotás során a rendvédelmi szervezetnek együtt kell működnie a társadalommal, a politikával és a gazdasággal. „A rendészet az államigazgatásnak az a területe, ami a leginkább ki van szolgáltatva a napi rögtönzéseknek, politikai csatározásnak és a legjobban nélkülözi a stratégiai szemléletet, a közszolgálat szakmaként való elismerését” (Finszter 2003: 13). A rendészeti stratégia alkotása tehát rendkívül fontos, mert meghatározza a rendőrség és a rendvédelmi

szervek munkájának irányát és prioritásait a közbiztonság fenntartásában és a bűnözés elleni küzdelemben. Ez lehetővé teszi, hogy a rendőrség hatékonyan koncentráljon azokra a területekre és bűncselekményekre, amelyek a legnagyobb fenyegetést jelentik a közösségre nézve. „Magyarországon 2014-ig egy a közigazgatáshoz közelített – ahol indokolt – civilisztett, centralizált (ugyanakkor bizonyos feladatok ellátását helyi/területi szintre decentralizáló/dekoncentráló), szolgálati ágai tekintetében integrált (monista), rendészetnek kell kimutatható eredményeket elérnie a közrend, közbiztonság és az államhatár védelmének megerősítésében, valamint a megelőző és fenyegetésekre való reagáló képesség fejlesztésében” (Dunavölgyi 2011: 19).

A stratégia kidolgozása segít meghatározni, hogy milyen erőforrásokra van szükség a célok eléréséhez és hogyan lehet azokat hatékonyan felhasználni. Ez segít elkerülni az erőforrások pazarlását és a felesleges kiadásokat. A rendészeti stratégia lehetővé teszi a proaktív megközelítést a bűnözés elleni küzdelemben. Az elemzések és a prediktív analitika segítségével a rendőrség képes előre jelezni a bűnözést és megelőző intézkedéseket hozni annak érdekében, hogy csökkentse a bűncselekmények számát.

A rendészeti stratégia kidolgozása során a lakosság is bevonható a folyamatba, melyhez a Bűnmegelőzés Nemzeti Stratégiájáról szóló 115/2003. OGY határozat ad magyarázatot az alábbiak szerint; „A modern európai felfogás szerint a közbiztonság ugyanis olyan kollektív társadalmi termék, amely az egyének és közösségeik tevékenységéből, az állami szervek hatósági intézkedéseiből, a polgárok önvédelmi képességei és a vállalkozói piac nyújtotta szolgáltatások együtteséből alakul ki. Ennek mérésére, elemzésére a fejlett demokráciákban egyre bonyolultabb és pontosabb módszereket alkalmaznak. Ilyenek például az áldozatok megkérdezésén vagy a lakosság széles köreiből végzett vizsgálatok eredményein alapuló elemzések. Nálunk ezek még nem részei az intézményesen szervezett tájékoztatásnak, ilyen típusú információk nem állnak a jogalkotó rendelkezésére.” A rendészeti stratégia segít meghatározni az adatgyűjtés és elemzés módszereit, amelyek révén a rendőrség képes követni az eredményeket és a stratégia hatékonyságát. Ez lehetővé teszi a stratégia finomhangolását és a folyamatos fejlesztést. A stratégia nyilvánosságra hozása és rendszeres beszámoló készítése a végrehajtásról növeli a rendőrség transzparenciáját és elszámoltathatóságát a lakosság és a döntéshozók előtt.

A modern MI rendészeti stratégia szükséglete

„A rendőrség egymagában nem képes közbiztonságot teremteni. A rendőrség közbiztonságot szervező munkájának értékelési dimenziói: az objektív, illetve a szubjektív közbiztonság. Az objektív közbiztonságot

szokás statisztikai mutatók alapján jellemezni. Objektív mérceként szolgálhat az elkövetett bűncselekmények, szabálysértések száma, a sértettek száma, az elkövetési érték alakulása. A rendőrséget jellemző számadatok arra alkalmasak, hogy a szervezet teljesítményét vizsgáljuk. A teljesítmény két minőségi kifejeződése, a hatékonyság és a hatásosság között éles határvonalat kell húznunk. Az eredményesség fogalmát nem tekintjük valós értelemben vett minőségi jelzőnek, mert csupán azt fejezi ki, hogy a rendőrség megtett valamit, ami a cselekvések számával kifejezhető” (Kis et al. 2014: 28). A stratégia kidolgozása és végrehajtása révén a rendőrség jobban felkészülhet a kihívásokra és a változó bűnügyi helyzetre, miközben növeli a lakosság szubjektív biztonságérzetét és a közbiztonságot.

Mint ahogyan arról a bevezetésben említést tettem, az Europol ugyancsak vizsgálja ChatGPT nyelvi modell rendészeti alkalmazhatóságát. A szervezet a csalást és a pszichológiai manipulációt, a dezinformációt és a kiberbűnözést három kritikus aggodalomra okot adó területként azonosította:

- *Csalás és pszichológiai manipuláció*
A ChatGPT azon képessége, hogy gyorsan generál valóság-hű szöveget, amely felhasználható gyanútlan személyek kihasználására. Ez a képesség felhasználható egy egyén vagy csoport beszédstílusának megszemélyesítésére, bizalmuk elnyerésére és pénzügyi vagy személyes veszteség helyzetébe való félrevezetésére.
- *Dezinformáció*
A ChatGPT rendkívül hatékony a szövegek gyors és nagy mennyiségben történő előállításában. Ez ideális eszközzé teszi a propaganda és a dezinformáció terjesztésére, mivel lehetővé teszi egy adott narratívát támogató üzenet gyors és hatékony előállítását és terjesztését.
- *Számítástechnikai bűnözés*
A ChatGPT nemcsak emberszerű nyelvet állít elő, hanem különböző programozási nyelveken is képes kódot generálni. Ez az alacsony szintű technikai tudás értékes eszközzé teszi a bűnözők számára, mert minimális erőfeszítéssel rosszindulatú kódokat tudnak előállítani és terjeszteni.

A jelentés azonosított néhány bűncselekményt a GPT-3.5 és a GPT-4 esetében is, megjegyezve, hogy néhány esetben a GPT-4 verzió potenciálisan káros válaszai még fejlettebbek voltak elődjénél. A ChatGPT gyors hozzáférést biztosít a felhasználóknak a bűncselekmények széles körével kapcsolatos információkhoz, a lakásbetöréstől a terrorizmuson át a gyermekek szexuális kizsákmányolásáig. Használható

rosszindulatú célokra, például adathalászatra, csalásra, személyi adatokkal való visszaélésre és bizalomra épülő manipulációra. Nagyon valóság-hű szövegeket generálhat a felhasználói kérés alapján, lehetővé téve a bűnözők számára, hogy meggyőző adathalász e-maileket és más megtévesztő dokumentumot, üzenetet hozzanak létre még korlátozott nyelvtudás mellett is. Emellett hamis közösségi média szerepvállalást is teremthet, hogy nagyobb legitimitást biztosítson a megtévesztő befektetési ajánlatoknak, valamint online dezinformációt, gyűlöletbeszédet és terrorista tartalmakat terjesszen. A ChatGPT lehetővé teszi a bűnözők számára, hogy gyorsan és egyszerűen hitelesen hangzó üzeneteket készítsenek, nagy mértékben megkönnyítve ezzel a potenciális áldozatok megtévesztését és bizalmuk elnyerését.

A ChatGPT jelenlegi verziója számos rosszindulatú célra képes alapvető eszközöket előállítani, így különösen hasznos azok számára, akik nem rendelkeznek kódolási vagy fejlesztési ismeretekkel. A ChatGPT legújabb verziója a GPT-4 jobban megérti a kód kontextusát, kijavítja a programozási hibákat. Ez a továbbfejlesztett verzió felbecsülhetetlen segítséget nyújt a kevés technikai ismerettel rendelkező potenciális bűnözőknek, és lehetőséget nyújt a haladó felhasználóknak a kifinomult számítógépes bűnözői technikák további finomítására és automatizálására. Az Europol ajánlásokat adott ki a bűnüldözési szervek számára, hogy felkészüljön a ChatGPT-hez hasonló nyelvtanuló gépekre, azaz a nagy nyelvi modellekre lehetséges következményeire. Ezek az ajánlások arra ösztönzik a bűnüldöző szerveket, hogy hívják fel a figyelmet az LLM-ek rosszindulatú használatának lehetséges káraitra, fejlesszék készségeiket annak megértésében, hogy az LLM-ek hogyan használhatók fel, valamint működjenek együtt az érdekelt felekkel annak biztosítása érdekében, hogy a biztonsági mechanizmusok prioritást élvezzenek e technológiák használatakor. Az Europol azt javasolja továbbá, hogy a bűnüldöző hatóságok vizsgálják meg a saját adataik alapján képzett, személyre szabott LLM-ek lehetőségét, feltéve, hogy tiszteletben tartják az alapvető jogokat (URL3). Míg az Europol munkaértekezletei a ChatGPT potenciálisan rosszindulatú felhasználási eseteinek azonosítására összpontosítottak, amelyek már ma is lehetségesek, a cél az is volt, hogy elemezzék ezeket a megállapításokat és ajánlásokat fogalmazzanak meg arra vonatkozóan, hogy a bűnüldöző szervek hogyan biztosíthatnak jobb felkészültséget az esetleges jövőbeli eseményekre. Megállapították, hogy a rendészeti szerveknek tisztában kell lennie a ChatGPT-hez hasonló LLM-ek pozitív és negatív következményeivel. A tudatosság elengedhetetlen a lehetséges joghézagok azonosításához és kezeléséhez, valamint a rosszindulatú használat megelőzéséhez. A bűnüldöző szerveknek meg

kell érteniük az LLM-ek hatását a különböző bűnözési területekre, hogy előre jelezzék és vizsgálják a technológiával való visszaélést. Fejleszteniük kell

továbbá a létrehozott tartalom pontosságának és esetleges torzításainak értékeléséhez szükséges készségeket (1. sz. ábra).

ERŐSSÉGEK	<ul style="list-style-type: none"> • Gyors információfeldolgozás: A ChatGPT képes nagy mennyiségű információt gyorsan feldolgozni és értelmezni. • 24/7 elérhetőség: A ChatGPT mindig elérhető, így a rendőrség bármikor használhatja. • Skálázhatóság: A rendőrségi szervek könnyen skálázhatják a ChatGPT rendszert a növekvő munkaterheléshez, anélkül, hogy nagyobb erőforrásokat kellene hozzárendelni. • Nyelvi képességek: A ChatGPT több nyelven is képes kommunikálni. 	GYENGESSÉGEK
FENYEGÍTETTÉSEK	<ul style="list-style-type: none"> • Adatvédelmi aggályok. • Félreértések: A ChatGPT által generált válaszok félreértéseket okozhatnak. • Hibás információ: Ha a ChatGPT hibás információt kap, akkor hibás választ is adhat. Ez különösen kritikus lehet a rendőrségi munka során. 	LEHETŐSÉGEK

1. sz. ábra: A ChatGPT nyelvi modell rendőrségi felhasználhatóságának SWOT analízise

Forrás: Saját szerkesztés

Az MI rendészeti stratégiának a rendőrség szempontjából az alábbi alapvető elemeket szükséges tartalmaznia:

- **Célkitűzések és stratégiai tervek**
Az MI rendészeti stratégiai céljai és a hosszú távú tervek lefektetése, beleértve a bűnügyek előrejelzését, a bűncselekmények megelőzését, a bűnüldözést, az adatvédelmet és a társadalmi felelősségtudatot.
- **Adatgyűjtés és adatkezelés**
Az adatgyűjtés és feldolgozás területein történő együttműködés az adatvédelmi előírások és etikai elvek betartásával (Pintér 2020).
- **Elemzés és értékelés**
Az adatok elemzése és modellezése az MI-rendszerben a bűnüldöző tevékenységek fejlesztése, jogsértések előrejelzése és megelőzése érdekében.
- **Minőségbiztosítás**
Az MI-rendszer teljesítményének értékelése, időszakos ellenőrzése, diagnosztizálása és javítása. Az MI-rendszer optimalizálása az elemzések és modellezések hatékonysága érdekében.
- **Képzés és fejlesztés**
Képzés az MI-rendszer karbantartására és működtetésére, közös munka az MI-rendszer

továbbfejlesztésében, valamint az együttműködési szint növelése más rendvédelmi szervezetekkel. Az MI-rendszer hatékonyabbá tétele a technológiai előrehaladás segítségével, újabb eszközök és eljárások bevezetése az MI-stratégia fejlesztése érdekében.

A hatékonyság növelése érdekében az MI-rendszereket össze kell hangolni a rendőrség célkitűzéseivel. A nyelvi modellek beépítését célzó rendészeti stratégián a rendészeti szervek által végzett tevékenységekbe a nyelvi modellek integrálási folyamatának tervezését értem. A nyelvi modellek – kiemelten a ChatGPT – rendészeti felhasználhatóságát célzó stratégiai pontokat a következő fejezetben mutatom be.

Magyarország MI stratégiájának rendvédelmi vetülete

Magyarország Mesterséges Intelligencia Stratégiája 2020-2030 a mesterséges intelligenciát akként határozza meg, „*mint a betáplált adatok alapján önmagukat tanítani és javítani képes algoritmikus rendszerek összessége*” (URL4). Az MI Stratégia 4.2.4 Államigazgatás – „Adatvezérelt szolgáltató állam” alfejezete alatt tesz javaslatot a Rendvédelmet szolgáló ellenőrzési rendszerek bevezetésére az alábbiak szerint:

- *Robotzaru program továbbfejlesztése*

- *Határvédelmi ellenőrző rendszerek fejlesztése, komplex azonosítási rendszer kialakítása*
- *Adataalapú, komplex elemzéseket használó bűnüldözés és bűnmegelőzés*
- *Meglévő MI technológiák bevezetése a nyomozási folyamatokba” (Magyarország Mesterséges Intelligencia Stratégiája 2020-2030).*

Az MI Stratégia mellékletében viszont intézkedési tervként kerülnek rögzítésre a rendvédelmet érintő főbb tárgykörök az alábbiak szerint:

- *“Robotzsaru program továbbfejlesztése*
- *Határvédelmi ellenőrző rendszerek fejlesztése, komplex azonosítási rendszer kialakítása és további fejlesztései*
- *Adataalapú, komplex elemzéseket használó bűnüldözés és bűnmegelőzés*
- *Támogató funkciók fejlesztése, és a szükséges kapcsolódó igazgatási, jogalkotási tevékenységek ellátása*
- *Bűnelkövetői kapcsolati hálózatok MI alapú feltérképezése, gráf adatbázis továbbfejlesztése*
- *Meglévő MI technológiák bevezetése a nyomozási folyamatba*
- *Szöveg leiratozás, szövegfelolvasás és nyelvtchnológiai funkcionalitás biztosítása és rendszerintegráció, a digitális iratok és ügyfélkapcsolatok során keletkezett rögzített beszéd elemzése szövegbányász eszközökkel, adatfeldolgozás*
- *Manuális adatkezelési, -feldolgozási folyamatok automatizálása*
- *Okmányokkal kapcsolatos adatkezelési és feldolgozási tevékenység támogatása*
- *Kompetenciafejlesztés, képzések, továbbképzések*
- *A szükséges infrastrukturális környezet biztosítása” (Magyarország Mesterséges Intelligencia Stratégiája 2020-2030).*

Az MI stratégia a részletes projektterv kidolgozásának határidejét 2021.03.31-re, míg a projekt végrehajtását 2024.03.31-re datálja. A következőkben az MI Stratégia mellékletében felsorolt szempontok vizsgálatát fogom végrehajtani.

A Robotzsaru program továbbfejlesztése

Számítógép használata a bűncselekmények analizálásakor először az 1960-as évek közepén történt meg St. Louisban, az USA-ban (Harries 1999: 6), amely nehézkesen induló motorként szolgált a bűnmegelőzésben és felderítésben. Ahogyan a szerző a címében utal a folyamatra: az út a „gombostűtől a számítógépig” tartott. A ChatGPT vagy más, hasonló MI alapú chatbotok integrálása a rendőrségi ügyfeldolgozó programokba sok előnnyel járhat, például az ügyek gyorsabb kezelésével és a rendőri erőforrások

hatékonyabb felhasználásával. Az alábbiakban néhány lépést és szempontot mutatok be az integrációhoz.

Jogi szabályozás és adatvédelem

„A legfontosabb előrelépés az MI átfogó szabályozása terén, hogy 2021 áprilisában a Bizottság közzétette az MI-ről szóló rendelettervezetere (Artificia Intelligence Act) tett javaslatát, amely fontos korlátozásokat tartalmaz az EU-ban vagy azzal összefüggésben használt MI-rendszerekre vonatkozóan” (Mezei 2023: 56). Az adatvédelem és a jogi szabályozás kulcsfontosságú szempontok, amikor a ChatGPT vagy bármilyen MI rendszert rendőrségi, vagy egyéb rendészeti, bünygyi programokban alkalmazunk. Az ilyen alkalmazásoknak szigorúan be kell tartaniuk az adatvédelmi és etikai elveket, valamint a jogszabályokat, hogy megvédjék az egyének jogait és személyes adatait. A rendőrségi alkalmazásoknak gondoskodniuk kell a felhasználók személyes adatainak biztonságáról és védelméről, tiszteletben kell tartania az egyének adatvédelmi jogait, beleértve a hozzáféréshez, javításhoz és törléshez való jogot. Ez magában foglalja a szigorú adatvédelmi intézkedések alkalmazását, például a titkosítást és a hozzáférési ellenőrzéseket. A rendőrségi alkalmazásoknak tisztázniuk kell, hogy milyen adatokat gyűjtenek, mennyi ideig tárolják őket és milyen célból használják fel. Az adatgyűjtésnek és tárolásnak összhangban kell lennie az adatvédelmi jogszabályokkal, melyre az Országos Rendőr-Főkapitányság (a továbbiakban: ORFK) jelenleg hatályos normatív szabályzói, így a Robotzsaru rendszer alkalmazásáról-támogatásáról, fejlesztéséről és használatához szükséges képzésről szóló 13/2023. (V. 31.) ORFK utasítás és az informatikai alkalmazásfejlesztési folyamatok szabályozásáról szóló 14/2022. (IV.7.) ORFK utasítás jelentenek garanciát.

Fontos, hogy az alkalmazások átláthatóak legyenek az adatok felhasználását illetően. A felhasználóknak tudniuk kell, hogy milyen adatokat gyűjtenek róluk és hogyan használják fel azokat. Az alkalmazásoknak kerülniük kell a diszkriminációt és az előítéleteket. Az MI rendszereknek objektíveknek és tisztességeseknek kell lenniük és nem szabad előnyben részesíteni vagy hátrányos megkülönböztetést tenni az alapján, hogy valaki például milyen etnikai csoportba tartozik. Minden rendőrségi alkalmazásnak be kell tartania az adatvédelmi és az emberi jogokat védő jogszabályokat. Az adatvédelmi és etikai kérdések kapcsán célszerű független szervezetek vagy szakértők részvételét is bevonni a felügyeletbe és az ellenőrzésbe, hogy biztosítsák az alkalmazások törvényességét és etikáját (Négyesi 2023). Az adatvédelem és a jogi szabályozás nagyon fontos az MI alapú rendőrségi alkalmazásoknál,

hogyan megvédjék az egyének jogait és az adatbiztonságot. Az ilyen rendszerek tervezése és alkalmazása során a technológiai fejlesztőknek és a jogalkotóknak szorosan együtt kell dolgozniuk a megfelelő etikai és jogi irányelvek kialakítása és betartása érdekében. A chatbotok felhasználóinak adatai gyakran nagyon érzékenyek lehetnek, mivel a rendőrség speciális hatósági jogosítványokkal rendelkezik, amelyek lehetővé teszik, hogy adatokat gyűjtsenek és használjanak fel.

Az alábbiakban bemutatom azon, számomra a legfontosabbnak tartott alapelveket, amelyek segíthetnek biztosítani a felhasználók biztonságát és adatvédelmét a rendőrségi chatbotok használata során:

- *Transzparencia és irányelvek*

Fontos, hogy a rendőrségi chatbot használata során a felhasználók tudják, milyen jogokkal rendelkeznek adataik védelme érdekében. A chatbotok használata előtt figyelmeztetni kell őket, hogy az adatok biztonsága és védelme érdekében a rendőrség milyen adatokat fog gyűjteni, milyen okból, valamint, hogy ezeket az adatokat milyen módszerekkel kezeli, tárolja. A rendőrségi chatbot használata során fontos a felhasználói adatvédelemhez kapcsolódó irányelvek előzetes deklarálása (Bicskei 2023: 106). Az adatvédelmi irányelvek tartalmazhatnak részletes információkat az adatok gyűjtéséről, azok felhasználásáról és azok szabályszerű kezeléséről. Fontos, hogy a rendőrség tájékoztassa ügyfeleit annak biztosítása érdekében, hogy tisztában legyenek ezekkel az irányelvekkel.

- *Személyes adatok védelme*

A rendőrségi chatbot használata során lényeges a személyes adatok védelme. Ha a chatbot személyes adatokat kér, akkor biztosítani kell, hogy azokat csak a törvényben meghatározott célból használja fel. A rendőrségről szóló 1994. évi XXXIV. törvény 77. §. alapján: „(1) A rendőrség feladatai ellátása során a) bűnmegelőzési és bűnüldözési (a továbbiakban együtt: bűnüldözési) feladatai ellátásához szükséges személyes adatokat, b) rendészeti és határrendészeti feladatainak ellátásához szükséges személyes adatokat, c) az általa lefolytatott közigazgatási eljárásban részt vevők személyes adatait, d) a segélyhívások fogadása keretében megismert személyes adatokat, e) a körözési eljárás lefolytatásához szükséges adatokat, f) a személyvédelmi és létesítménybiztosítási feladatai ellátásához

szükséges adatokat kezeli. Az adatvédelmi elveket a legjobb gyakorlatokkal összhangban biztosítsa a rendőrség, amelyek a hatósági irányelvek betartásával együtt alkalmazhatók.

- *Biztonsági szabályok*

A chatbot alkalmazásával, a felhasználók védelme érdekében lehetőség van a biztonság fejlesztésére. Az adatok károsításának elkerülése végett a rendőrségnek biztosítani kell az ellenőrzést és az optimalizálást azáltal, hogy a chatbot rendelkezik a korszerű biztonsági intézkedésekkel, illetve a folyamatos felügyeletet is biztosítja. A rendőrségnek előzetesen ki kell dolgoznia egy olyan biztonsági stratégiát is, amely minimális kockázattal rendelkezik és hosszabb távon biztosítani tudja az adatok védelmét (Necz 2022).

Az integráció technikai kivitelezése

Az alkalmazás programozási interfésze, azaz az Application Programming Interface (a továbbiakban: API) integrációja a ChatGPT vagy más MI rendszerekkel egy rendőrségi programban azért lehet hasznos, mert lehetővé teszi a rendőrök és a bűnüldöző hatóságok számára, hogy gyorsan és hatékonyan kérjenek információt, vagy alkalmazzanak MI-alapú eszközöket munkájuk során. A ChatGPT API-ján keresztül a fejlesztők képesek integrálni a nyelvi modellt más alkalmazásba vagy egyéb szoftverbe (URL5). Az API lehetővé teszi a fejlesztők számára, hogy hozzáférjenek a ChatGPT és Whisper modellekhez, amelyek révén lehetővé válik a nyelvi feldolgozás és a beszéd-felismerés. Az ChatGPT fejlesztőjének weboldalán, az OpenAI platformon található leírás tartalmazza a modellek használatát és azok informatikai paramétereit (URL6).

A ChatGPT API használatához több lehetőség is rendelkezésre áll. A hivatkozások áttanulmányozását követően látható, hogy az integrációt megelőzően ki kell választani egy olyan API-t, amely a kívánt MI szolgáltatást kínálja. Például OpenAI GPT-3 API-t lehet használni a ChatGPT integrálásához. Az adott rendőrségi program fejlesztőjének elő kell készíteni a szükséges fejlesztési környezetet és eszközöket az API integrációhoz. Regisztrálni kell az API szolgáltatónál és hozzáférést kell szerezni az API-hoz. Ez általában egy API kulcs használatát jelenti, amelyet az integrált alkalmazás használ az API-val történő kommunikációhoz. Ki kell alakítani egy felhasználói interfészt, amely lehetővé teszi a szerverek számára az MI rendszerrel való kommunikációt. Ez lehet egy webes

felület, egy mobilalkalmazás vagy egy parancssori felület, attól függően, hogy hogyan tervezték az alkalmazást.

A jelenlegi Web rendszerű RZS NOVA felületek erre alkalmasabbnak látszanak, mint a telepítést igénylő Robotzsaru NEO Integrált Ügyviteli és Ügyfeldolgozó rendszer (a továbbiakban: Robotzsaru NEO) alkalmazás. Természetesen az Robotzsaru NEO jelenleg is tartalmaz több olyan hivatkozást, ahol gyakorlatilag a webböngészőn keresztül nyit meg alkalmazásokat, például a NOVA Vezetői Modul, a NOVA TIR egységkezelő vagy akár a fegyverzeti modul.

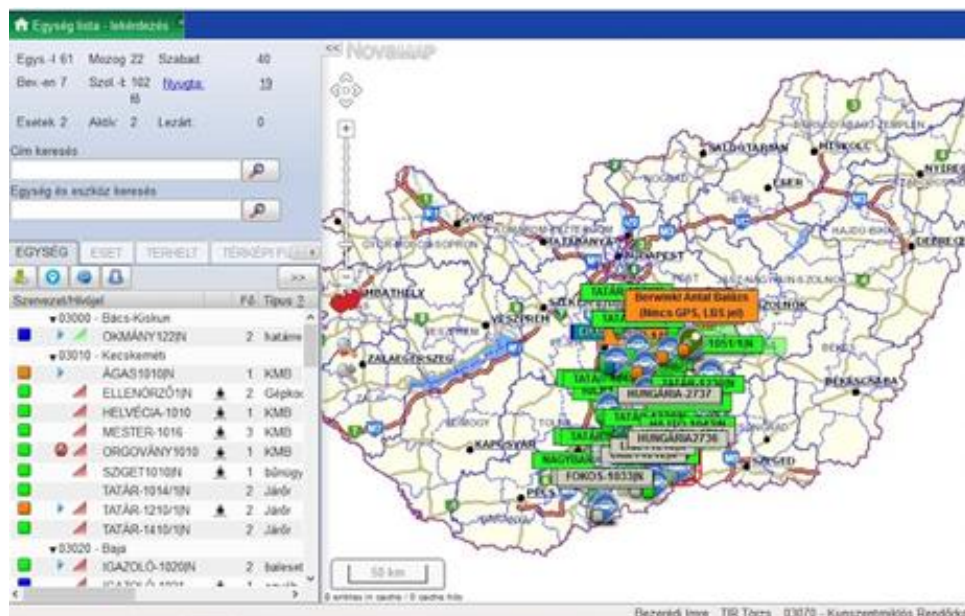
Az API használatokor is kiemelten kell gondoskodni az adatvédelmi és biztonsági intézkedésekről. Az integrációt alaposan ki kell próbálni és folyamatosan tesztelni kell, hogy biztosított legyen annak megfelelő működése és stabilitása. Az API integrációt rendszeresen kell karbantartani és frissíteni, hogy az mindig a legfrissebb verziót használja és megfeleljen a változó jogi és biztonsági követelményeknek. Az API integrációval a rendőrségi programok hatékonyabban és gyorsabban szerezhetnek hozzáférést az MI szolgáltatásokhoz, amelyek segíthetik a rendőrség munkáját.

Használati esetek meghatározása

Az 1980-as, 90-es évek között felgyorsult technikai fejlődésnek köszönhetően folyamatosan kerültek bevitelre rendészeti adatok számítástechnikai

eszközökbe, így térinformatikai adatok megjelenítésére is sor került. Erre szükség is volt, ugyanis a XX. század második felében nemzetközi szinten folyamatosan nőtt a bűncselekmények száma (Borbíró 2009), ami nem volt másként hazánkban sem. A rendszerváltáshoz közeledve a felgyorsult bűnözési hullám újabbnál-újabb módszerekkel ostromolta a társadalmat, majd az ezredfordulót követően sorra jelentek meg az elkövetői módszerek között informatikai eszközök alkalmazása.

A rendőrség ma már nap mint nap dolgozik térinformatika segítségével. Kezdetben, a 2000-es évek elején a Robotzsaru NEO rendszerben kerültek GEO kódolásra bűncselekmények elkövetési helyek szerint, azok büntetőjogi minősítése mellett, azonban a rögzített adatok nem voltak hitelesek. Ennek számtalan oka volt, ám hiányoztak az adatrögzítés elvének általános szabályai, a megfelelő informatikai rendszerek rendelkezésre állása, a szükséges térkép szoftverek megléte. Ezt követően, pár évvel ezelőtt a szolgálattervezés és szolgálatirányítás professzionalizálása mentén vezették be a rendőri egységek térinformatikai rendszerbe történő bevitelét, melyhez hozzájárultak a szolgálati gépjárművekben, majd kézi eszközökben elhelyezett GPS eszközök is. A Tevékenység-irányító Központok 2013. évi bevezetésével a rendőrségi vezetés jól látható képet kapott a szolgálatot teljesítő erőkkel kapcsolatban, amely már egy új webes alapon futó úgynevezett Robotzsaru NOVA rendszerben működő térképrendszerben kapott helyet (2. sz. ábra).



2. sz. ábra: A Rendőrség NOVA TIR térképrendszerének működése

Forrás: Saját szerkesztés

A Robotzsaru NOVA térképészeti rendszere előrelépés a Robotzsaru NEO térképészeti alrendszeréhez képest, itt ugyanis különböző térkép rétegek között lehet választani, így a különböző katonai és európai térképek mellett határjeleket és egyéb POI-kat is lehet alkalmazni. A Robotzsaru NOVA helyt ad mindemellett a csapatszolgálati törzsfeladatok ellátásának dokumentálására, vagy akár különböző személyi kompetenciák rögzítésére (pl. bombakutató, vezetői engedély típus, nyelvtudás), melyek az egységek kezelésében előnyhöz juttatják az irányító parancsnokokat. Pusztán egy lépésre vagyunk a sikertől, azaz informatikai megoldást találni a két rendszer között, hogy a Robotzsaru NEO-ban rögzített GEO kódok, a Robotzsaru NOVA térképészeti alrendszerében a különböző szűrők – jogsértés ütköztetése, ezáltal szabálysértés, bűncselekmény akár egyidejű megjelenítésére sor kerülhet, időpont, szervezeti elem, vagy település – segítségével implementálásra kerüljenek.

A Robotzsaru integrált ügyviteli, ügyfeldolgozó és elektronikus iratkezelő rendszer egységes és kötelező használatáról, jogosultsági rendjéről, az adatvédelem, valamint a rendszerfejlesztés előírásairól szóló 18/2011. (IX. 23.) ORFK utasítás² 42. pontja már releváns adatok közé sorolta a GEO kódokat (*a Magyarországon elkövetett azon bűncselekmények földrajzi koordinátái, amelyek helyszínhez köthetők*), majd a 43. pont meghatározta, hogy az ügyfeldolgozás, a rendőri intézkedések rögzítése során az ügy előadója köteles a releváns adatokat rögzíteni és karbantartani. Az utasítás ezzel megteremtette a GEO kódok, mint releváns adatok kezelésének kötelezvényét is. „A bűnözésföldrajzi kutatások eredményeinek, azon belül kiemelten a bűnözési térképek adaptálására számos lehetőség kínálkozik a gyakorlati rendőri munka során. Felhasználhatók – többek között – a bűnmegelőzés, a szolgálatsszervezés, az elemző-értékelő munka, valamint a rendészeti vezetés és oktatás területén. Emellett a bűnözési térképek hasznosak lehetnek a területfejlesztési koncepciók, városrendezési dokumentumok kidolgozásánál is” (Mátyás–Sallai 2014: 335). A rendszerváltást követően a rendőrség átalakulására, megreformálására a várakozásokkal ellentétben nem került sor, azonban a különböző kormányok egységes politikai konszenzus és ezáltal rendészeti stratégia megalkotása helyett saját kormányzásuk időszakában folyamatos átszervezésekkel, modernizációs folyamatokkal kívánták egyensúlyozni a rendőrség bűnözésre adott válaszreakciós képességét. Ez nagyban befolyásolta a szervezet proaktív szerepének

kialakulását, pontosabban annak elmaradását. A ChatGPT Robotzsaruba integrálása számos használati esetet hozhat létre, amelyek segíthetik a rendőrök munkáját és a bűnüldözést. A rendőrök információt kérhetnek az MI rendszertől a bűnözéssel vagy gyanús tevékenységekkel kapcsolatban. Az MI segíthet a bűnügyi adatok gyors keresésében és elemzésében, lehetőséget teremthet a bűnügyi adatok elemzésében és értékelésében, például a bűnügyi statisztikák vagy az elkövetői mintázatok alapján. Ez segítheti a bűnüldöző hatóságokat a bűnözés megértésében és ennek megelőzésében. Az MI rendszer segíthet a nyomozás során a szöveges információk elemzésében, például gyanús üzenetek vagy tanúk vallomásainak értelmezésében. A ChatGPT online platformokba vagy mobilalkalmazásokba – mint például a Rutin³ – történő integrálásával válaszolhat a megkeresésekre vagy egyéb általános bejelentésekre, illetve magyarázatot adhat egyes rendőrségi intézkedésre. Naprakész információval szolgálhat egyes folyamatban lévő nyílt nyomozásról, helyi bűnözési trendekről, újonnan megjelenő fenyegetésekről. Az MI rendszer segíthet a rendőröknek a jogi kérdések gyorsabb és pontosabb megválaszolásában, például a rendőri eljárásokkal vagy a jogi előírásokkal kapcsolatban. A rendszer alkalmazható lehet válsághelyzetekben, például az eltűnt személyek felkutatásában vagy veszélyhelyzetek kezelésében. A használati eseteket az adott rendőrségi program céljai és igényei alapján kell meghatározni, majd az integrációt ennek megfelelően kell tervezni és kialakítani. Az MI rendszerrel való együttműködés lehetőségei szinte korlátlanok, a technológia folyamatosan fejlődik, amely támogatni tudja a bűnüldöző hatóságok munkáját.

Egyéb chatbot integrálási lehetőségek

Mint korábban bemutattam, jelenleg a webes felület a legelterjedtebb és könnyen elérhető platform a felhasználók számára. Ezen keresztül a rendőrök és a lakosság egyaránt hozzáférhetnének a chatbot szolgáltatásaihoz böngészőjük segítségével. Egy mobilalkalmazás kényelmes lehetőséget kínál mind a rendőröknek, mind a lakosságnak a chatbot elérésére. Ez különösen hasznos lehet a közterületen dolgozó rendőröknek. Ilyen mobilalkalmazás a már létező Robotzsaru TIR Mobil, amely Android alapú rendőrségi ügykezelő, ügyfeldolgozó és főként prioráló, valamint intézkedést kezelő applikáció. Ezen applikáció alkalmas lehet arra, hogy a ChatGPT integrálását követően számos, az előző alcímben felsorolt funkciót

² Hatályát veszítette 2023.10.01-től

³A RUTIN egy, az úton lévőket segítő rendőrségi applikáció, amely új verziója a Google Play és App Store online áruházakból ingyenesen letölthető. (Androidos eszközökhöz már most is elérhető,

iOS platformra hamarosan érkezik a frissített app.) A közlekedőket arra biztatjuk, hogy próbálják ki és használják az okostelefon-alkalmazást.” <https://www.police.hu/hu/hirek-es-informaciok/utinfo/rutin> [Letöltve: 2023.09.28.]

működtesse. Révén mobil applikációról van szó, az alkalmazó rendőr néhány adat rögzítését követően képes lenne jelentést, jegyzőkönyvet készíteni, egy-egy deliktum helyszínén ok-okozati összefüggéseket alaposabban felmérni, az elkövetőt kézre keríteni.

Kalmár egy 2016-ban készített tanulmányában (Kalmár 2018) már rámutatott a TIR Mobil applikáció automatizált jelentéskészítő funkciójának lehetőségére, melynek megvalósításához a ChatGPT ez idáig hiánypótló volt. Léteznek azonban rendőrségi fejlesztéstől független⁴ applikációk és weboldalak is, melyek ugyancsak a főként a közterületen intézkedő rendőrök tevékenységét segítik. Egyes rendőrségi alkalmazásoknál a parancssori interfész is lehetőséget kínálhat a chatbot használatára, például rendőri számítógépek vagy rendszerek integrálásakor. A hangvezérlés lehetőséget biztosít a rendőröknek, hogy verbálisan kommunikáljanak a chatbot-tal különösen akkor, ha kéz nélkül szeretnék használni az alkalmazást, mint például vezetés vagy intézkedés közben. Fontos lehetőség a chatbot bevezetése az SMS vagy telefonos hívások útján történő kommunikációra, különösen olyan helyeken, ahol a mobilinternet hozzáférés korlátozott lehet. Az alkalmazásoknak általában többféle felületet kell támogatniuk annak érdekében, hogy minél szélesebb körű hozzáférést biztosítsanak a rendőröknek és a lakosságnak. A választott felületnek egyszerűnek és intuitívnek kell lennie, valamint meg kell felelnie a felhasználók igényeinek és elvárásainak.

Határvédelmi ellenőrző rendszerek fejlesztése, komplex azonosítási rendszer kialakítása és további fejlesztései

Az MI és a chatbotok jelentős hozzáadott értéket és hatékonyságot hozhatnak a határvédelmi rendszerekbe. Az MI képes automatikusan észlelni és azonosítani a jogsértéseket, amelyek felmerülhetnek a határon, mint például az embercsempészet, az illegális bevándorlás, vagy egyéb illegális csempész tevékenységek. Az MI képes folyamatosan figyelni az érzékelőket és a kamerákat, azonosítani a gyanús tevékenységeket, és figyelmeztetni a kezelő szervet. Képes előrejelzéseket készíteni a határon várható eseményekről a korábbi adatok és információk alapján. Ez lehetővé teszi a hatóságok számára, hogy proaktív lépéseket tegyenek a potenciális kockázatok kezelése érdekében. Az MI képes

automatikusan döntéseket hozni a határon, mint például az átjárók vagy kapuk nyitására, zárására, figyelmeztetések küldésére vagy egyéb intézkedésekre. Ez jelentősen csökkentheti a rendőri erő szükségességét és az emberi hibákból eredő kockázatokat.

Személyazonosítás és azonosítási ellenőrzés

A határvédelemben a személyazonosítás fontos tényező a biztonságos határellenőrzés érdekében. A ki- és belépő személyek azonosítása biztosítja a hatóságok számára, hogy tudják kik lépnek be az országba, valamint megakadályozzák a veszélyes személyek belépését a területre. Az azonosítási folyamat során a hatóságok a beutazó személyek személyazonossági igazolványát, útlevelet, vízumkérelmét, valamint a fotó- és ujjlenyomatokat ellenőrzik. A határrendészet számára fontos, hogy megfelelő képzésben részesüljenek a rendőrök és határrendészek, amely során elsajátítják az azonosítási és ellenőrzési gyakorlatot, valamint a megfelelő technikák és eszközök felhasználását. Az azonosítási eljárások betartása és a megfelelő eszközök felhasználása hozzájárul a határvédelem hatékonyságához, biztonságához és elősegíti az országok polgárainak védelmét. *„A folyamatos forgalomnövekedés és az illegális bevándorlás fokozódó veszélye manapság jelentős nyomást gyakorol a határátkelőhelyekre, és az ott szolgálatot teljesítő személyzetre. Ezenkívül a hosszadalmas ellenőrzési procedúrákból fakadó lassú határátlépés befolyásolja az utazók elégedettségét és a kereskedelmet is. Megoldásként a határátkelőhelyeken megkezdődött a „mesterséges intelligenciával támogatott önkiszolgáló ügyintézési oszlopok”, más néven KIOSK-ok, és a hozzájuk tartozó automata átléptető kapuk bevezetése. Az önkiszolgáló automata KIOSK-os rendszerek alkalmazása a modern határátléptetési folyamat kötelező eleme, amely a személyes adatok biztonságos kezelése mellett megkönnyíti és felgyorsítja a határátlépési folyamatokat a résztvevők számára”* (Hertelendi – Hornyik 2022: 211).

A rendőrség számos vizuális információ-feldolgozó eszközt használ munkája során. Az arcfelismerés a vizuális információ-feldolgozó eszközök egyike. A határrendészeti arcfelismerés az arcfelismerés technológiájának egy alkalmazási területe, amelyet határőrségi és határrendészeti szervek használnak azonosítási és biztonsági célokra a határok ellenőrzése során. Az arcfelismerés egy biometrikus azonosítási módszer, amely az egyén arcának különböző jellemzőit használja azonosításra vagy azonosítási célokra (Necz

határszakaszának teljes hosszában 11.472 fő rendőri erőre lenne szükség ahhoz, hogy pusztán a határterületen történő egyvonalas alkalmazás működőképes legyen (BALLA–KUI 2017: 230)

⁴ Ilyen weboldal a jelenleg fejlesztés alatt álló Wisecop is. <https://wisecop.webnode.hu/> [Letöltve: 2023.09.12.]

⁵ Említést érdemel az a számítás, amelyet Balla József és Kui László végeztek. Ennek során arra jutottak, hogy Magyarország 1 103 kilométeres schengeni külső

2020). A rendőrség határrendészeti feladatának ellátása során már használja az arcfelismerést azonosításra és ellenőrzésre. Az arcfelismerés segít beazonosítani azokat a személyeket, akik illegálisan lépnek be a schengeni térségbe, vagy akik hamis vagy hamisított okmányokkal próbálnak átjutni a határon. Amikor valaki átlépi a határt, az arcfelismerő rendszerek képesek összehasonlítani a személy arcát az adatbázisokban található képekkel, például útlevélképekkel vagy korábbi felvételekkel, hogy megerősítsék az azonosítást vagy ellenőrizzék az egyén személyazonosságát (Nagy 2023). Az adatbázisokban található képek alapján az arcfelismerő rendszerek figyelmeztethetik a rendőröket, határvadászokat, ha olyan személyek lépnek be az országba, akikre jelzés van elhelyezve a Schengeni Információs rendszerben (Dobó–Gyaraki 2021). Az arcfelismerés hasznos lehet az irreguláris migráció (Ritecz–Sallai 2016) megfigyelésében is. Amikor valaki illegálisan lép be egy országba vagy területre, az arcfelismerő rendszerek segíthetnek azonosítani és nyomon követni az ilyen személyeket, ehhez azonban nem elegendő az MI-t a magyar rendőrség alkalmazásaihoz hozzárendelni, hanem célszerű lenne azt legalább Schengeni szinten összehangolni.

Kommunikáció és támogatás

A nyelvi modellek képesek lehetnek megérteni és kommunikálni a különböző nyelveket beszélő külföldi személyekkel, migránsokkal, illetve elemezni a szöveges- és hang alapú információkat, amelyek kapcsolódnak a határrendészeti feladatokhoz. A chatbotok, mint a ChatGPT használhatóak a határvédelmi rendszerek információszolgáltatására és a személyekkel való kommunikációra. Az emberek kérdéseire válaszolva és tájékoztatást nyújtva segíthetnek a határvédelmi folyamatok megértésében és az eljárásokkal kapcsolatos kérdések megválaszolásában. Az AI hatékonyabbá teheti a határvédelmet és lehetővé teszi a hatóságok számára, hogy proaktívan kezeljék a határon felmerülő kockázatokat, miközben egyszerűsítik a határellenőrzési folyamatokat.

Adatalapú, komplex elemzéseket használó bűnüldözés és bűnmegelőzés

A ChatGPT nagy mennyiségű adatot képes gyorsan és hatékonyan feldolgozni. Ez a funkció különösen előnyös a parancsnoki állomány és az elemző-értékelők számára. A ChatGPT elemzi mind a számokat, mind a szavak adatkészleteit, és írásos összefoglalást készít az adatokról. „Az MI-rendszer legfőbb előnye, hogy az emberhez képest nagyságrendileg több adatot tud beszerezni és összevetni, addig a digitális nyomozószoftver esetén már nem

az adatok nagy mennyiségű előállítás, hanem éppen azok csökkentése, szűkítése jelenti a nagy előnyt” (Herke 2021: 1720). A Chat GPT a kódszkriptekben is találhat problémákat és jártas a Python és a Java nyelvben.

Adatgyűjtés és elemzés

„A nyílt forrásból származó információgyűjtés nem alacsonyabb értékű, vagy kevésbé fontos, mint a titkos vagy bizalmas forrásból származó információszerzés” (Gulyás 2023: 1242). Az MI képes konkrét adatgyűjtésre a rendelkezésére álló adatmennyiségből. Előfordul azonban számos eset, amikor a ChatGPT még „hallucinál” (Kis 2023). A ChatGPT esetében a hallucinálás gyakran arra utal, hogy a rendszer olyan válaszokat generál vagy információkat ad, amelyek nem felelnek meg a valóságos tényeknek vagy nem helytállóak. Ilyen lehet például téves információk vagy összefüggéstelen válaszok generálása. A ChatGPT ilyen hallucinációi általában azért fordulhatnak elő, mert a rendszer nem rendelkezik valóságon alapuló érzékeléssel vagy tudatossággal, és az inputjaira reagál anélkül, hogy tudna róla azok mennyire helyesek vagy valóságosak. Ezért fontos, hogy bármilyen információt vagy választ, amit a ChatGPT vagy más MI rendszer generál, kritikusan értékeljék és ne tekintsek feltétlenül megbízhatónak. Az emberi felhasználóknak mindig saját értékítéletükre kell hagyatkozniuk és ellenőrizniük kell az információkat, különösen olyan kritikus vagy érzékeny területeken, mint a biztonság, az egészségügy vagy a jog. A ChatGPT segíthet az adatok gyűjtésében és elemzésében a bűncselekmények, a bűnözési minták és az események nyomonkövetése érdekében. Nagy mennyiségű adatot gyorsan fel tud dolgozni és elemzi is azokat. Ez lehetővé teszi a rendőrség számára, hogy gyorsan azonosítsa az elkövetőt és hatékony elemzéseket készítsen bűnözési mintákról, ill. egyéb fontos információkról.

Chatbotok és lakossági információszolgáltatás

A chatbotok integrálása lehetővé tenné a rendőrség számára, hogy gyorsan és hatékonyan válaszoljon a lakosság kérdéseire, miközben csökkentené az ügyfeldolgozó folyamatok terhet. A rendőrségi chatbot legalkalmasabb felülete attól függene, hogy milyen célokat és feladatokat kívánna elérni vele a rendőri szerv. Az MI rendszer segíthet automatikus válaszokat generálni olyan gyakran ismétlődő kérdésekre, mint például az elérhetőségi információk vagy a lakoságnak szóló tájékoztatók. A chatbotok, mint a ChatGPT használhatóak lakossági információszolgáltatásra és kapcsolattartásra. Az MI alapú chatbotok segíthetnek a lakosság kérdéseinek és panaszainak kezelésében, valamint a lakoságnak nyújtott információk

terjesztésében. A chatbotoknak pontosan kell érteniük az ügyfelek kérdéseit és pontos válaszokkal kell szolgálniuk azokra. Emellett fontos, hogy könnyen elérhetőek és könnyen használhatóak legyenek. Csak azokat az információkat kell megosztaniuk, amelyek mindenki számára elérhetőek (mint például a jogszabályok), vagy amelyeket a rendőrség külső kommunikációnak szánt (például sajtóhírek). Nem szabad személyes adatokat vagy szolgálati titkot közölniük. A chatbotok azonnal válaszolhatnak a kérdésekre és segíthetnek az információgyűjtésben. Képesek lehetnek a helyi bűnügyi statisztikák és információk közvetítésére a lakosok számára. Ez segíthet az embereknek abban, hogy tisztában legyenek a környékükön zajló bűnügyekkel és óvintézkedéseket tehessenek, ezzel segítve a bűnmegelőzést. A ChatGPT tanácsokat adhat arról, hogyan lehet az otthonokat biztonságosabbá tenni vagy hogyan lehet elkerülni a veszélyeket. A chatbotok azonban nem helyettesíthetik a humán kapcsolattartást és párbeszédet, a rendőrségnek továbbra is szorosan együtt kell működnie a lakossággal, hogy hatékonyan biztosítsa a közrendet és a közbiztonságot.

A chatbotok és a kriminálstratégia

A rendelkezésre álló magyar szakirodalom többnyire a következő német-osztrák irányú fogalmat használja: „A kriminálstratégia a kriminalisztika azon része, amely a bűnözés helyzetére vonatkozó objektív adatokra és a lakosság szubjektív biztonságérzetére alapozva a bűnözés elleni küzdelmet szolgáló intézkedések összességének tervezésével és szervezésével foglalkozik. Ennek során figyelembe veszi a szervezeti célt meghatározó jogi, gazdasági, kulturális, történelmi, politikai és szociális keretfeltételeket” (Berthel–Lapp 2017: 27). Ezen fogalom értelmezése során látható, hogy a német definíció a kriminalisztika, mint a bűnüldözés tudománytana alá helyezi a kriminálstratégiát. Teszi ezt a magyar irodalomban Fenyvesi Csaba is, aki a kriminálstratégiát a kriminalisztika alatt a krimináltechnika, krimináltaktika és kriminálmetodika mellett együttesen nevesíti (Fenyvesi 2004: 757). Finszter Géza (2018) nyelvi és terminológiai okok miatt helyesebbnek látja a bűnüldözési stratégia elnevezés használatát. Finszter szerint a kriminálstratégiát ugyanis helyesebb bűnüldözési stratégiának nevezni: „A bűnüldözési stratégia (az idegen nyelvű szakirodalomban kriminálstratégia) magában foglalja a bűnüldözés célkitűzéseit, valamint – e célok megvalósítása érdekében – a bűncselekmények elleni küzdelem legeredményesebb eszközeinek és módszereinek a távlati fejlesztését” (Finszter 2018: 86). Finszter egy másik irodalomban kifejti, hogy kriminálstratégia alatt a rendőrségi erők

kriminálpolitikai célokat szolgáló tervező tevékenységét érti, amely átfogja a szervezet kialakítását, a személyi és tárgyi feltételek folyamatos biztosítását, valamint a sikerhez fűződő együttműködés megteremtését. A kriminálstratégiát nem szűkíti le pusztán felszámoló jellegre, megfogalmazásában a represszív és preventív rendészeti metódust egyformán, ötvözve nevesíti (Lévay 2016: 803).

A kriminálstratégia segít abban, hogy a rendőrök megfelelően reagáljanak a bűnözésre és a közbiztonság fenyegetéseire. Elősegíti a hatékony stratégiai gondolkodást és döntéshozatalt. Előirányozza az elemezni szükséges adatokat, az azonosítandó bűnügyi trendeket és meghatározza a főbb prioritásokat a hatékony és eredményes rendőri fellépés érdekében. A ChatGPT használata lehetőséget biztosít a rendőrségnek, hogy hatékonyabban megértsék és elemezzék a bűncselekményekkel kapcsolatos adatokat. Azok az információk, amelyeket a ChatGPT eleméz és értékel, segíthetnek az adatok rendszerezésében, a trendek és minták azonosításában, valamint a bűncselekmények megelőzése és felszámolása terén hasznos stratégiák tervezésében. A ChatGPT képes összegyűjteni és elemezni a lakossági bejelentéseket és észrevételeket. Ezek az információk segíthetnek a rendőrségnek az erőforrások stratégiai elosztásában.

Rendszeres értékelés és visszajelzés

Fontos, hogy a beépített MI rendszereket rendszeresen értékeljék és finomítsák a hatékonyság és az eredmények javítása érdekében. A rendszeres visszajelzések segíthetnek az MI rendszerek fejlesztésében és optimalizálásában. A ChatGPT és rendőrség együttes rendszerének értékelése és visszajelzése rendkívül fontos ahhoz, hogy folyamatosan javítsák a rendszer működését és hatékonyságát. Az értékelésnek és a visszajelzésnek egyaránt tartalmaznia kell a rendszer teljesítményét, a felhasználói élményt és a hasznosított adatok minőségét. A felhasználói visszajelzések és a rendszer teljesítményét értékelő adatok segítenek azonosítani a hiányosságokat és az ígéretes területeket, valamint javaslatokat adni azokra a fejlesztésekre, amelyek növelhetik a rendszer értékét. Összességében a rendszeres értékelés és visszajelzés kulcsfontosságú ahhoz, hogy folyamatosan javítsuk a ChatGPT és a rendőrségi rendszer hatékonyságát, és hogy biztosítsuk a felhasználóbarát és biztonságos rendőrségi tevékenységet.

Támogató funkciók fejlesztése és a szükséges kapcsolódó igazgatási, jogalkotási tevékenységek ellátása

A rendőrség és a ChatGPT-t támogató funkcióinak fejlesztése hozzájárulhat a hatékonyabb és eredményesebb rendőrségi tevékenységhez. Az alábbiakban felsorolunk néhány olyan lehetséges funkciót, amelyek fokozhatják a rendszer hatékonyságát:

- *Chatbot automatizálása*
A chatbot rendszer automatizálása lehetővé teszi, hogy a rendőrségi ügynökök több időt töltsenek a közvetlen emberi interakcióval és a chatbotok kezelhessék a kisebb feladatokat, például javaslatok vagy egy korai vonalas standard eljárások terén.
- *Természetes nyelvfeldolgozás (a továbbiakban: NLP)*
Az NLP-funkciók a ChatGPT és a rendőrségi rendszer között egyértelműbb kommunikációt lehetővé tevő funkciók, amelyek automatikusan megértik, hogy mi az ügyfél igénye és problémája, majd gyors választ adnak a kérdésekre (Deng–Lin 2022).
- *Felkészítési adatelemzés*
Az adatelemzés hasznos az információk megértéséhez és az interakciók értékeléséhez, amelyek javíthatják az ügyfélélményt és lehetővé teszik a rendőrség számára, hogy könnyebben kezelje a felmerülő kihívásokat.
- *Hangvezérlés*
A hangvezérlés lehetőséget nyújt a rendőrségi állománynak a hangparancsok aktiválására az ügyféllel való személyes interakció során.
Ez csak néhány példa a ChatGPT és a rendőrség szempontjából hasznos funkciókra, amelyek fejlesztését érdemes megfontolni tevékenységük támogatása és hatékonyságuk növelése érdekében. A fejlesztés lakossági visszajelzéseket igényel, hogy a felhasználókat érintő és releváns funkciók jobban érvényesüljenek a gyakorlatban. A jogalkotás csak lohol a technológia után (URL7), és az adatvédelemmel kapcsolatos aggályok miatt több országban is betiltották a ChatGPT használatát. Az olasz adatvédelmi hatóság például átmenetileg tiltotta be a ChatGPT használatát Olaszországban, mivel az OpenAI sérti az adatvédelmi szabályokat a személyes adatok gyűjtésére vonatkozó

gyakorlatával. Az EU-s szabályok nem engedik meg, hogy az MI korlátlanul, bármilyen adatból tanuljon. A lehetséges kockázatok csökkentése érdekében a ChatGPT használatát megelőzően kiemelten fontos ügyelni arra, hogy az adatvédelmi szempontok már a tervezési fázisban a projekt részévé váljanak, beépüljenek és megfelelő hangsúly kerüljön többek között az adatkezelési célok világos kijelölésére, a lehetséges adatbiztonsági kockázatok azonosítására és kezelésére.

Bűnelkövetői kapcsolati hálózatok MI alapú feltérképezése, gráf adatbázis továbbfejlesztése

A bűnelkövetői kapcsolati hálózatok MI alapú feltérképezése és a gráf adatbázis továbbfejlesztése rendkívül fontos az eredményesebb bűnüldözés érdekében. Az MI alapú feltérképezés lehetővé teszi a bűnözői kapcsolatok gyors és hatékony azonosítását, ami segíti a rendőrséget a bűnügyek gyorsabb felderítésében és megelőzésében. Az MI alapú feltérképezés során a rendőrségnek speciális adatokat kell összegyűjtenie, többek között telefonos beszélgetéseket, e-maileket, közösségi média posztokat, GPS nyomkövető adatokat és banki tranzakciókat. Az adatok elemzése és feldolgozása során az MI algoritmusok a kapcsolati hálózatokat detektálják és ábrázolják, amelyek megmutatják az egyes bűnözők közötti kapcsolatokat és a bűncselekmények szervezetségének szintjét. A gráf adatbázis továbbfejlesztése pedig lehetővé teszi, hogy a rendőrök hatékonyan vizualizálják a bűnözői kapcsolatokat és gyorsan megismerjék azokat a gyűjtött adatokat, amelyek segíthetnek az ügyek megoldásában. A gráf adatbázisok arra összpontosítanak, hogy a kapcsolatokat és az azok közötti összefüggéseket jól strukturálják, ennek eredményeként egyedi megjelenítést és elemzést tesznek lehetővé. Az adatok megjelenítése gráfok segítségével segít a rendőröknek az emberek vagy események közötti kapcsolatok megértésében, a hálózatok diagramjának használatával pedig lehetővé teszi a rendőrség számára, hogy az adatokat könnyebben összehasonlítsa és azonosítsa az egyes szervezetekben végzett tevékenységek és bűnpártolások alapján. Az MI alapú feltérképezés és a gráf adatbázis továbbfejlesztése együttesen segítheti a rendőrséget a bűnözői kapcsolatok hatékonyabb feltérképezésében.

Meglévő MI technológiák bevezetése a nyomozási folyamatba

A meglévő MI technológiák bevezetése a nyomozási folyamatba nagyobb hatékonyságot és gyorsaságot eredményezhet a rendőrség számára. Mint ahogyan az előzőekben ismertettem, az MI technológiákat használhatják adatok elemzésére, adatgyűjtésre, bűnözői kapcsolatok feltárására, illetve a rendőrök és bűnüldöző hatóságok hatékonyabb támogatására. Az MI technológiákat azonban nem csak az adatelemzésben, hanem az adatgyűjtésben, például a szemtanúk kihallgatásának feldolgozásában is használhatják. Az MI alapú fordítóprogramok lehetővé teszik a rendőrségi nyelvészeti szakemberek számára az idegen nyelvű kihallgatások gyorsabb és egyszerűbb végrehajtását és értelmezését. A ChatGPT segíthet egyes nyomozati cselekmény felgyorsításában is, mint például egy szembesítés vagy bizonyítási kísérlet. Figyelmeztetheti az eljáró nyomozót, ha kihagyott egy jogi figyelmeztetést, avagy elmulasztott egy fontos eljárási aktust. A ChatGPT olyan adatok elemzésére is képes, mint a DNS-minták és a szöveges üzenetek, amelyek segítségével bonyolult jogi ügyekben is segíthet. A ChatGPT képes arra is, hogy bűnözői profilokat hozzon létre és előrejelzéseket készítsen a bűncselekményekkel kapcsolatban. Az MI technológiák a rendőrségi felszerelésekben és berendezésekben is felhasználhatók, például a drónokban, amelyek a rendőröknek segítenek a távoli területeken sokkal hatékonyabban történő észlelésben és azonnali reagálás biztosításában (Déri 2022). Az MI alapú technológiák bevezetése a nyomozási folyamatba biztosítani kívánja, hogy a rendőrség biztonságosabban dolgozzon és nagyobb hatékonysággal folytasson eljárást a bűncselekmények elkövetői ellen.

Automatizált adatgyűjtés

A ChatGPT automatizált rendőrségi adatgyűjtése egy olyan innovatív megközelítés lehet, amely hozzájárulhat a hatékonyabb és gyorsabb rendőrségi tevékenységhez. A ChatGPT rendszer a természetes nyelvfeldolgozás technológiájára épül (Deng–Lin 2022), célja, hogy automatikusan értelmezze és feldolgozza az ügyfelek által adott információkat, javaslatokat és kéréseket. Ezt a rendszert az adatgyűjtésre is felhasználhatjuk oly módon, hogy a rendőrségi nyomozásokhoz kapcsolódó adatokat automatikusan gyűjti be. Az NLP technológiájának használatával a ChatGPT rendszer a rendőrség számára nem csak azonnali értesítéseket ad, hanem élő adatokat is szolgáltat. Az automatizált rendőrségi adatgyűjtés és az NLP technológiájának használata segítheti a rendőrséget abban, hogy időt takarítson meg, valamint

gyorsabb és hatékonyabb tevékenységet folytasson a bűnözőkkel kapcsolatos információk és nyomozások gyűjtésében, feldolgozásában. Fontos azonban, hogy a nyelvi modell algoritmusai jelenleg tiltja a személyes adatok gyűjtését, ezért a chatbot számára az automatizált adatgyűjtés elsődlegesen olyan monitoring tevékenységet jelentene, melynek során a rendőrség által előzetesen megadott tárgyköröket kutatná (Bagó 2023: 207).

Prediktív rendészeti vetületek

„A rendőrség soha nem volt a geopolitika változója, de az AI miatt az lesz. A prediktív rendőrségek – mint amilyen a *Minority Report* [Különnvélemény] című filmben szereplő *PreCrime* – már az előtt jelzik a bűncselekményeket, hogy azok megtörténnének, s be tudják azonosítani az elkövetőket is. Bizonyos embereket célba is vehetnek – épp ezért a legveszélyesebb benne a módszer, amely alapján az AI eldönti, hogy ki a bűnöző, s ki nem. Ez az állampolgárok manipulálásának és befolyásolásának egy egészen új szintjét, dimenzióját vetíti előre, amit egyfajta új típusú „technológiai gyarmatosításnak” is nevezhetünk. Ez oda vezethet, hogy egyes országokban minden ok nélkül külföldieket tartóztathatnak le, vagy egy-egy külpolitikai cél elérése érdekében manipulálhatják a prediktív rendőrségi rendszereket. Ezért a szerző szerint lehet, hogy a jövőben azok az országok lesznek a legbiztonságosabbak, ahol nem működik majd ilyen rendőrség, mivel akkor az általa okozott konfliktusok sem jelentkeznek” (Sütő 2023: 147).

A nemzetközi rendészettudományi szakirodalom az adatok hasznosíthatóságára vonatkozó kutatásokat és azok gyakorlati alkalmazását a „predictive policing” gyűjtőfogalom alatt tárgyalja (Szabó 2019: 20). A prediktív rendészet olyan megközelítés, amely a rendőrségi tevékenységet adatelemzésre és előrejelző elemzésre építi. Ennek a hatékony, adatvezérelt rendészetnek célja az, hogy megakadályozza a bűncselekményeket, mielőtt azok bekövetkeznek. A prediktív rendészet előrejelzést ad arról, hogy egy adott helyszín vagy terület milyen valószínűséggel lesz érintett bűnözéssel. Adatok elemzésén alapul, amelyeket az eltérő bűnözési minták azonosítására használnak fel. Ez lehetővé teszi a hatóságok számára, hogy az időben és a helyszínre összpontosítva, proaktív módon lépjenek fel a bűnözés megelőzésének érdekében. A prediktív rendészethez használt chatbotok új dimenziót vezethetnek be a bűncselekmények előrejelzésében és megelőzésében. A chatbotok nagyon hatékonyan képesek tanulni és előrejelezni a felhasználó által adott információk alapján (Kuhn 2023). A chatbotok használatával a rendőrségi szervek részére könnyedén rendelkezésre állhatnak a különböző adatforrások, mint például a bűnözéssel kapcsolatos információk, és be tudják építeni ezeket a chatbotokba előrejelzési

alkalmazás céljára. Az ilyen chatbotok lehetővé teszik a lakosság számára, hogy maguk jelentsék be a környezetükben bekövetkezett gyanús eseményeket, körülményeket. Ez segíthet a bűnözés megelőzésében, és egyúttal biztosítja az adatbiztonságot, ill. a privát szférát. Számos algoritmust használnak a bűncselekmények elkövetésének megakadályozására (előre jelezve a lehetséges elkövetési helyeket vagy elkövetőket). Idetartozhat az úgynevezett hot-spot elemzés (magasabb bűnözési veszéllyel rendelkező földrajzi helyek meghatározása, például RTM: Risk Terrain Modeling) és a bűnözés összekapcsolása (crime linking, ami múltbeli bűncselekmények összevetése bizonyos – már azonosított vagy még azonosítatlan – személyekkel annak érdekében, hogy megjósolják, hol és mikor követik el a következő bűncselekményt, például HART: Harm Assessment Risk Tool) (Herke 2023).

A hot-spot úttörőjeként tartják nyilván a Kansas-i kísérletet. „... Amelynek során az amerikai várost két egyenlő szektorra osztották, akként, hogy az egyik részben megduplázták, a másikban pedig felére csökkentették a rendőri jelenlétet. Ezután hosszabb időn keresztül mérték a közbiztonság alakulását a két területen. Azt találták, hogy a jogsértések száma nem emelkedett ott, ahol a korábban megszokotthoz képest fele annyi rendőr teljesített szolgálatot, viszont nem javult a helyzet abban a körzetben, ahol a rendőrök száma megduplázódott. Kész volt a diagnózis: a rendőri jelenlét semmiféle befolyással nem bír a közbiztonságra... Az Egyesült Államok egyik városában, Minneapolisban 1998-ban megismételték a Kansas City-i kísérletet, de merőben más előkészítéssel. Rendőrségi szakemberek sok tapasztalat birtokában jelölték ki a város úgynevezett „forró pontjait”, amelyek különösen fertőzettek voltak a jogsértésektől. Ezután ezekben a körzetekben megduplázták, más helyeken viszont megfelezték a rendőri jelenlétet. A közbiztonság számottevően javult, egy év alatt a büntetendő cselekmények száma jelentősen csökkent. Bebizonyosodott, hogy amennyiben a közbiztonsági jelenlét megszervezésénél a szakmai jártasságra építünk, és amennyiben alkalmazkodunk a lokális szükségletekhez, az eredmény nem marad el” (Finszter 2012: 133).

A prediktív rendészet részben tehát az adatelemzés és a gépi tanulási algoritmusok használatának módszere azon területek azonosítására, ahol a bűncselekmények a legvalószínűbbek. Ez a megközelítés azon az elgondoláson alapul, hogy a bűnözés nem véletlenszerű. Inkább azt bizonyítja, hogy a bűnözésnek vannak mintái, amelyek azonosíthatók, rajzolhatók és megjósolhatók. Aggályok merülnek fel azonban az algoritmusokban rejlő elfogultság és megkülönböztetés lehetőségével kapcsolatban. Ez azért van, mert a ChatGPT-t az emberek egy hatalmas adatkészleten keresztül táplálják. Fennáll továbbá a hamis jelzések kockázata, emiatt téves küldések, hot-spotok születnének bizonyos területen. „A digitális tér csupán

lehetőséget nyújt a nyomozó hatóság számára. A siker záloga a digitális adatoknak a kriminalisztikai látásmódon alapuló, a kiforrott nyomozási módszertannal („kriminalisztikai ajánlásokkal”) kombinált, ebből fakadóan értő és hatékony feldolgozása” (Bogdány 2018: 180). Az MI képes lehet azonosítani a bűncselekményeket és a veszélyes helyzeteket a közösségi médiában és az online kommunikációkban. Ez lehetővé teszi a hatóságoknak, hogy gyorsan cselekedjenek és reagáljanak a veszélyhelyzetekre.

Szöveg leiratozás, szövegfeldolvasás és nyelvtechnológiai funkcionalitás biztosítása és rendszerintegráció, a digitális iratok és ügyfélkapcsolatok során keletkezett rögzített beszéd elemzése szövegbányász eszközökkel, adatfeldolgozás

Az MI és a ChatGPT alkalmazása a rendészeti szöveg leiratozásában, szövegfeldolvasásban és nyelvtechnológiai funkcionalitás biztosításában jelentős előnyöket hozhat a rendőrségi munka hatékonyságának növelésében. Az MI technológiák automatizálása és gépi tanulása hozzájárulhatnak a rendőrségi jelentések hatékonyabb és pontosabb elkészítéséhez. Az MI alapú eszközök segítségével a rendőrök gyorsabban és hatékonyabban írhatnak jelentéseket. Az NLP technológiája lehetővé teszi például a gépi tanulási algoritmusoknak, hogy megértsék a természetes nyelven írt szövegeket és automatikusan feldolgozzák azokat (Németh–Virágh 2023). Az MI rendszerek képesek megtanulni az adott rendőri szervezet jelentéseinek mintáit és stílusát, így képesek javaslatokat tenni a rendőröknek a szövegek szerkesztése során. Az MI technológiák segíthetnek továbbá az információk automatikus adatbázisba vagy rendszerbe való importálásában és elemzésében is. Fontos azonban megjegyezni, hogy az MI technológiák nem szabad, hogy teljesen kiváltsák a rendőrök szerepét a jelentésírásban. A szakmai tapasztalat és az emberi empátia, a diszkrecionális jog gyakorlása továbbra is elengedhetetlen a releváns és megbízható jelentések készítéséhez. „Csakhogy az is kiderült, hogy a ChatGPT már az empátiát igénylő helyzetekben is – olyanokban például, amikor rossz híreket kell közölni a beteggel – jobb mondatokat képes generálni, mint az ember” (Bialkó 2023). Az MI technológia csak egy eszköz, amely segíthet a rendőröknek a hatékonyabb és pontosabb jelentések elkészítésében.

Szöveg leiratozás

Az MI segíthet a rendőrségi nyomozás során keletkezett rögzített beszéd (például hangfelvételek

vagy telefonbeszélgetések) szöveggé alakításában. Az automatizált szövegleiratozás lehetővé teszi a beszélgetések könnyebb és gyorsabb elemzését. A ChatGPT kiváló munkát végez a háttérzajok kiszűrésében és több ember hanganyagának elkülönítésében. Még lenyűgözőbb az idegen nyelvek fordításának képessége. Ha egy rendőr rögzít egy interakciót és a zajok vagy egyéb körülmények miatt nem érti, ami abban elhangzik, az MI révén dekódolhatja és leszűrheti egyes személyekre. A Chat GPT videóelemző eszközei még fejlesztés alatt állnak, de fennáll annak lehetősége, hogy a rendszer egy személy biometrikus adatait használja, majd ezeket a biometrikus adatokat használja ugyanazon személy megtalálására egy másik videóforrásban. Csak idő kérdése, hogy ez a lehetőség mikor lesz elérhető.

Szövegfelolvasás

Az MI képes lehet szövegek hangos felolvasására, amely segíthet a rendőrség állományának és nyomozóknak az iratok és információk könnyebb hozzáférésében, különösen akkor, ha nincs lehetőség a szövegek hagyományos elolvasására. A ChatGPT lehetővé teszi a virtuális asszisztensek, chatbotok és az automatikus szövegfelolvasók fejlesztését. A rendőrség az automata szövegfelolvasókat használhatja a nyomozásokhoz szükséges információk megszerzéséhez nélkülözhetetlen adatok gyűjtésére és értelmezésére. A szövegfelolvasók segítségével a rendőrség könnyen szűrheti az online platformokon a bűncselekményekkel kapcsolatos beszélgetéseket, támadásokat, fenyegetéseket és az illegális tevékenységeket. Az MI alapú szövegfelolvasók javítják az események minőségét és pontosságát, így a rendőrség hatékonyabban és pontosabban tájékozódhat a nyomozási folyamatban. Az MI technológia segítségével a rendőrség a szövegfelismerési és adatelemzési technikákat alkalmazhatja az online platformok szöveges tartalmának teljes skálájára.

Nyelvtechnológiai funkcionalitás

Ide tartozik a nyelvi elemzés, a gépi fordítás, a nyelvészeti elemzések és az automatizált nyelvértelmezés. Az MI és a nyelvtechnológiai funkciók lehetővé teszik, hogy a rendőrség hatékonyabban és gyorsabban dolgozzon a szöveges adatokkal. Az MI és a nyelvtechnológiai funkciók segíthetnek a rendőrségnek a bűncselekmények szöveges dokumentációjában, a tanúvallomások és a bizonyítékok elemzésében, a nyomozások nyomon követésében, az adatok feldolgozásában és a rendszerszintű magyarázatok

készítésében a bűncselekmények elkövetésének módjáról és az azokkal kapcsolatos összefüggések elemzéséről. A nyelvtechnológia hatékonyan képes automatizálni az adatok feldolgozását, szűrését és rendszerezését. Az MI és nyelvtechnológiai funkciók alkalmazása emellett lehetővé teszi azt is, hogy a rendőrség képes legyen a bűnügyi nyomozásokhoz szükséges információkat kézzel írott jegyzetek, beszélgetések és audio felvételek formájában hatékonyan kezelni.

Adatfeldolgozás és szövegbányászat

A chatbotok képesek tanulni és előrejelzéseket készíteni a felhasználó által megadott információk alapján, különösen akkor, ha párhuzamosan vannak integrálva a különböző adatforrásokból származó adatkínálatokkal. A szövegbányászat a kapcsolódó szöveges információk feldolgozására szolgál, amelyek a bűnügyi mintákkal és az értesítésekkel kapcsolatosak. A szövegbányászati módszerek segítenek az információk automatikus elemzésében és feldolgozásában, mellyel az eseményekre világosabb képet lehet kapni. Az eszközök lehetővé teszik a rendőrség számára, hogy a kapcsolódó adatokat olyan formára hozzák, amelyekben egyszerűen is lehet keresni, így a rendőrség gyorsan reagálhat a bűnözők, a bűncselekmények és az érintett területek által meghatározott trendekre. Az ilyen adatfeldolgozási technikák és szövegbányászati módszerek nagymértékben javítják a rendőrségi munkát. Az adatok elemzése alapján a rendőrség gyorsan feltérképezheti a trendeket és felmérheti a reagálási lehetőségeket. Az adatelemzésen alapuló rendőrségi munka nagymértékben javíthatja a közrendet és a biztonságot, ugyanakkor megőrizheti az emberi jogokat és a polgárok privát szféráját is.

Manuális adatkezelési, -feldolgozási folyamatok automatizálása

A rendőrségi chatbotok segíthetnek az adatkezelési és adatfeldolgozási folyamatok automatizálásában, különösen azokban az esetekben, amikor a rendőrség állománya problémába ütközik az adatok feldolgozásával, azok rendszerezésével vagy kezelésével. A chatbotok képesek gyorsan és hatékonyan dolgozni és könnyen hozzáférhetővé teszik a rendőrségi adatbázisokat. Például, ha a 112-es segélykérőn⁶ a hívásfogadóknak sok beérkező hívása van, amelyeket kezelni és feldolgozni kell, a chatbot segíthet automatizálni a folyamatot. Kiszűrheti a téves hívásokat, így kevesebb teher nehezedne a humán erőforrásra. A chatbotok a rendőrségi adatbázisokkal is

⁶ Lásd még: <https://www.police.hu/hu/112>

integrálhatóak, így könnyen kereshetnek az adatbázisban található információk között és kinyerhetik a szükséges adatokat azokból. Emellett segíthetnek az adatok rendszerezésében, például a bűnügyi jelentések szervezésében és kategorizálásában. A chatbotok képesek jelentős időmegtakarítást nyújtani, csökkentve a rendőrség állományának terheit, mivel nem kell manuálisan megkeresni az információkat, hanem automatizált folyamatban dolgozhatnak ezekkel. Az adatok kezelésének és feldolgozásának automatizálása segít az adatok minőségének javításában és hatékonyságában is.

Okmányokkal kapcsolatos adatkezelési és feldolgozási tevékenység támogatása

Az MI, például a ChatGPT alkalmazása a rendészeti okmányokkal kapcsolatos adatkezelési és feldolgozási tevékenység támogatására számos módon történhet. Az MI képes lehet a digitális okmányok, például személyi igazolványok, útlevelek vagy hatósági jelzések automatizált kezelésére és azok digitalizálására. Az automatizált okmányfelismerés és az adatok kinyerése segíthet a rendőrségnek az okmányok azonosításában és az adatok gyors elérésében. A digitális okmánykezelés szintén kritikus szerepet játszik a rendőrség munkájában, mivel az elektronikus dokumentumokat és adatokat kell megszerezni és archiválni. A digitális okmánykezelés lehetővé teszi a rendőrség számára az információk osztályozását, kategorizálását és tárolását az elektronikus formában érkező információk kezelésekor. Ez hatékonyabb és hibamentesebb dokumentumkezelést tesz lehetővé, emellett jelentősen csökkenti a papírhasználatot és javítja a dokumentumok elérhetőségét és megosztását (Sallam 2023). Az MI képes nagy adatmennyiségeket feldolgozni és elemzéseket végezni, így a digitális okmányellenőrzéseknél a rendőrség gyorsan és hatékonyan tudja elemezni az okmányok adatait, például az okmányokban szereplő személyi adatokat, aláírásokat és más azonosító információkat. Az MI technológiák és az önjáró megoldások lehetővé teszik a digitális okmánykezelés automatizálását és hatékonyabbá tételét. Ezáltal a rendőrség hatékonyabban tudja kezelni és karbantartani az elektronikus dokumentumokat, ill. az adatokat könnyebben elő lehet hívni és elemezni.

Kompetenciafejlesztés, képzések, továbbképzések

„Az elmúlt körülbelül száz évben a világ mezőgazdaságiból előbb ipari, majd szolgáltató társadalommá alakult át, az oktatási rendszernek elég ideje volt alkalmazkodni és átalakítani a módszereit. A mostani változások pár évtized

alatt borítják a feje tetejére a munkaerőpiacot és az oktatás reakcióideje igen lassú” (Segal, 2018: 132). A ChatGPT alkalmazása a rendészeti területen a kompetenciafejlesztés, a képzések és továbbképzések terén is fontos lehet. Az MI és a gépi tanulás (a továbbiakban: ML) technológiák folyamatosan fejlődnek (Kollár 2018), és a rendőrségi állománynak képesnek kell lennie azok hatékony alkalmazására és kezelésére. Az alábbiakban bemutatok néhány módszert arra, hogy hogyan lehet az MI-t integrálni a rendészeti kompetenciafejlesztési és képzési tevékenységekbe.

Az MI segíthet az interaktív és testreszabott oktatási anyagok létrehozásában, amelyek segítenek az állománynak megérteni az MI és ML technológiák működését és alkalmazását (Demeter–Mező 2023: 78). A rendőrségi chatbotok segíthetnek a szervezeteknek kompetenciafejlesztéssel is. Az ilyen chatbotok a rendőrségi ügyekkel és folyamatokkal kapcsolatos információk és szolgáltatások széles spektrumát kínálják, amelyeket a rendőrök használhatnak, hogy javítsák és növeljék készségeiket. A chatbotok által kínált rendőrségi információk és szolgáltatások segíthetnek a rendőröknek az egyes tények és trendek jobb megértésében, valamint azok hatékony kezelésében. A chatbotok olyan eszközként is használhatóak, amelyek segítik a rendőrök oktatását és képzését. Az olyan fontos készségekkel is foglalkozhatnak, mint az interperszonális kommunikáció, a helyzetmegoldó képesség, a jogi ismeretek és a bűnüldözési stratégiák. A chatbotok emellett olyan szolgáltatásokat is kínálnak, amelyek segítenek az új alkalmazottak beilleszkedésében és az alapvető kompetenciák fejlesztésében. Ezek a szolgáltatások magukban foglalják a rendőrségi szakzsargont, a szervezeti működéssel kapcsolatos információkat és az ügyfélkezelési készségeket. Az ilyen chatbotok abban is segíthetnek, hogy a rendőrségi szervezeteket kevésbé terheljék az alapvető képzési és oktatási tevékenységek, mivel a chatbotok a különféle feladatokat „feladat-alapú” képzésekként és oktatási funkciókként is elláthatják. Ez azt jelenti, hogy a chatbotok széles körű elméleti ismereteket és gyakorlati tapasztalatokat is biztosítanak a rendőröknek, ám nem azonosítják a hiányos tudásokat és készségeket, illetve nem használják a rendőrök teljesítménye alapján az értékelési eljárásokat. Ennek előnye, hogy a rendőrség munkálatai folyamán emberek értékeljék az adott rendőr teljesítményét. A rendőrségi chatbotok használata a rendőrség munkájára fordított időn keresztül segíti a rendőrök fejlődését, valamint biztosítja az optimális megfelelést a beosztás miatt támasztott elvárásoknak, mindezt a meglévő rendőrségi rendszerbe könnyen integrálva.

Szimulációs tréningek

A chatbotok használata a rendőrségi szimulációs tréningekkel is együttműködhet. A rendőrség szimulációs tréningeket használhat, hogy a rendőröket több támadás- elhárítási mintára is felkészítse, mivel a tréningek különböző helyzetekben tesztelik a rendőrök reakcióit, szkennelik a környezetet és a helyes döntéseket, amelyeket aktívan ösztönöznek a rendőrök számára (Bagó 2023: 198). A chatbotok olyan eszközként is alkalmazhatók az éles helyzetekre való felkészülésben, amelyek segítik a rendőröket az éles helyzetekhez kapcsolódó információk előzetes megismerésében, az események megértésében, illetve a problémás helyzetek megoldásában, amelyek nehezen vagy ritkán fordulnak elő. A rendőrségi chatbotok és szimulációs tréningek együttműködése a rendőrség, valamint az egyének és csoportok számára egyszerűbbé és hatékonyabbá teszi a fejlődést.

Online kurzusok

A rendőrségi online kurzusok jó lehetőséget jelentenek a rendőrök számára, hogy saját tempójukban és helyüktől függetlenül sajátíthassák el azokat az információkat és készségeket, amelyek szükségesek ahhoz, hogy munkájukkal kapcsolatos előrehaladást érjenek el. A chatbotok a rendőrségi online kurzusokon keresztül képesek előzetes kérdések és válaszok, interaktív feladatok és rövid videók formájában segíteni a rendőröket a témák megértésében. A chatbotok segítségével a rendőrök interaktívan képesek tanulni, ami elősegíti az információ és a készségek jobb értését, ill. megértését. A chatbotok a szakmai képzésekben nagy tapasztalattal rendelkező szakértők tudását sajátítják el, amely az online kurzusok keretében hatékonyabbá teszi a tanulást. Az ilyen chatbotoknak is lehetősége van arra, hogy visszajelzést adjanak a rendőröknek, amely segít eldönteni, hogy van-e szükségük arra, hogy az online kurzusra a sztetend tanfolyamokon kívül további anyagokat is beiktassanak. A chatbotok a rendőrségi online kurzusokon keresztül a rendőrökkel való kommunikációra is alkalmasak, ennek köszönhetően könnyen elérhetőek és használhatóak akár az online kurzus keretében, akár azon kívül is.

Mentorálás és tanácsadás

A mentorálás célja, hogy az újonnan belépő vagy tapasztalt rendőrök számára támogatást, útmutatást, tanácsot és lehetőséget biztosítson a fejlődésre és a jobb teljesítményre. A chatbotok segíthetnek az újonnan belépő rendőröknek a különböző kérdések megválaszolásában, amelyek felmerülnek a rendőrségi munkájuk során. Gyakran van kérdésük, és még nem

biztosak abban, hogy kitől kell megkérdezniük, avagy nem merik feltenni kérdésüket. A chatbotok segíthetnek abban, hogy a rendőrök megbirkózzanak azzal a stresszel, amely szakmai tevékenységükhöz kapcsolódik. A chatbotok elérhetősege folyamatos, amely segíthet a rendőröknek, amikor szakmai szükségletét érzik annak, hogy szót váltsanak egy tapasztalt munkatársukkal.

A szükséges infrastrukturális környezet biztosítása

A ChatGPT vagy más MI rendszerek rendőrségi használatához megfelelő infrastrukturális környezetet kell biztosítani. A ChatGPT használata és integrálása a rendőrségi infrastruktúrába jelentős felelősséget ró a rendőrségre (Gyaraki 2023), hogy biztosítsa az infrastruktúra biztonságát és megbízhatóságát a chatbotok és virtuális asszisztensek számára. A ChatGPT-t gyakran az internetre csatlakoztathatják, így fontos, hogy a rendőrség biztosítsa az infrastruktúrát a biztonságos és megbízható internetkapcsolatok létrehozására. Biztosítani kell a megfelelő tűzfalakat, a vírusirtó programokat és malware védelmet is a chatbotok és virtuális asszisztensek komplex rendszerei elleni védelem érdekében. A rendőrségnek különös figyelmet kell szentelnie az adatvédelemre, ill. a tárolási szabványoknak megfelelően kell kezelnie az adatvédelmi kérdéseket. Biztosítani kell a megfelelő adatvédelmi politikákat, biztonsági szabványokat és eljárásokat a ChatGPT-k és virtuális asszisztensek számára. Az infrastruktúrának képesnek kell lennie a chatbotok és virtuális asszisztensek további kiegészítéseire, újra programozására és karbantartására, hogy azok a rendőrség által kezelt adatokkal és rendszerekkel való kölcsönhatás során hatékonyan működjenek.

Szerverek és számítógépek

Az MI rendszerek futtatásához erős szerverek és számítógépek szükségesek, amelyek képesek a komplex számítási feladatok végrehajtására és a modell tréningjére.

Adattárolás: az MI rendszerek nagy mennyiségű adattal dolgoznak, így szükség van megfelelő adattárolási infrastruktúrára. Ezt lehetőség szerint skálázható adattároló rendszerekkel kell megoldani.

Hálózati kapacitás

A rendőrségi chatbot használata jelentős hálózati kapacitást igényelhet. A chatbot működéséhez szükséges internetkapcsolat és a szükséges hardver- és szoftverinfrastruktúra költséges lehet, és a chatbot által forgalmazott adatokhoz szükséges hálózati kapacitás

biztosítása szintén jelentős költségeket jelenthet. A hálózati kapacitás jelentős szerepet játszik a rendőrségi chatbotok működésében. A fenti módszerek használatával lehetőség van arra, hogy a rendőrség megfelelően használja hálózati kapacitását és biztosítja a felhasználóknak a kellemes és hatékony chatbot élményt (Farkas 2023).

Monitorozás és karbantartás

A rendőrségi chatbotok monitorozása és karbantartása elengedhetetlen a botok sikeres működéséhez és hosszabb távú fenntarthatóságához. A chatbotok karbantartása magában foglalja a hibák javítását és az új funkciók hozzáadását, míg a monitorozás a botok performance elemzésére és a felhasználói interakciók értékelésére irányul. A chatbotok karbantartása során folyamatosan figyelemmel kell kísérni a botok működését, és azonnal javítani kell az esetleges hibákat. A botok karbantartása során a tesztek elvégzése, az automatizált tesztrendszer felülvizsgálata és az emberi tesztek elvégzése kulcsfontosságúak lehetnek ahhoz, hogy az alkalmazás megfelelően működjön. A chatbotok működésének monitorozása segíthet abban, hogy gyorsan reagáljanak és javítsanak a botokban lévő hibákra vagy megfelelően kezeljék a felhasználók kérdéseit. Az alkalmazás működése és a felhasználói interakciók elemzése lehetővé teszi a rendőrségnek, hogy jobban megértse a felhasználói igényeket és megfelelően reagáljon rájuk. Az ilyen elemzések alapján lehet javaslatot tenni az alkalmazás további fejlesztéseire és az esetleges problémák megoldására. A chatbotok biztonsága fontos szempont a karbantartás során. A rendőrségi chatbotokat rendszeresen kell auditálni annak érdekében, hogy az adatváltozások nyomon követhetők legyenek és a chatbot védett legyen az esetleges támadásoktól.

Képzés és szakértelem

A rendőrségi chatbotok kiépítése és fejlesztése jelentős erőforrást és szakértelmet igényel. A chatbotoknak meg kell érteniük a felhasználói megkereséseket. A chatbotok fejlesztése és karbantartása nagyfokú technikai tudást igényel a programozási nyelvek és adatbázisok ismeretében. A programozási adottságok javítása a karbantartási folyamatok során alapvető fontosságú. A rendőrségi chatbotoknak számos olyan kérdésre kell válaszolniuk, amelyeket a felhasználók feltehetnek. Éppen ezért fontos, hogy a fejlesztők megértsék a felhasználói szükségleteket és azt, hogyan lehet a botok által kínált tartalmat jobban igazítani a felhasználói igényekhez. A chatbotok fejlesztőinek ismerniük kell a legújabb

adatvédelmi törvényeket és követelményeket, illetve biztosítaniuk kell a rendőrségi felhasználók személyes adatainak biztonságát. A chatbotok fejlesztése során a fejlesztő csapatoknak ismerniük kell az aktuális trendeket és technológiákat. Az ilyen ismeretek lehetővé teszik az alkalmazások hatékonyabbá tételét a felhasználók számára és új funkciók bevezetését az alkalmazásokba. A rendőrségi chatbotok fejlesztéséhez szükséges szakértelem széles spektruma magában foglalja a nyelvtechnológiát, az adatelemzést, a programozást, a felhasználói interakciókat, az adatbiztonságot és az új trendek és technológiák ismeretét. A rendőri szervezetnek különösen fontos, hogy olyan szakemberekkel dolgozzon, akik képesek az alkalmazásokat a felhasználói igényeknek megfelelően kialakítani és használatba állítani.

Összegzés

Az Europol mellett számos külföldi és magyar kutató foglalkozott már az MI, valamint a nagy nyelvi modellek rendészeti vetületével. Törekedtem ezen részeredményeket komplex, elsősorban a magyar rendőrség által hasznosítható javaslatra rendszerezni a Magyarország Mesterséges Intelligencia Stratégiája 2020-2030 kormányzati dokumentumban foglalt és lefektetett rendvédelmet szolgáló ellenőrzési rendszerek nyomán. A rendőrség által jelenleg használt Robotzsaru NEO program továbbfejlesztése kapcsán a felhasználók és az adatkezelésben érintettek biztonságának érdekében tettem jogi és etikai javaslatokat. Bemutattam a ChatGPT integráció technikai kivitelezésének alapjait, ill. ajánlásokat tettem a nagy nyelvi modell és egyes rendőrségi alkalmazások, valamint eszközök együttes használatára, mint például a NOVA TIR térképészeti alrendszerének fejlesztésére, többek között a bűnügyi vagy baleseti adatok elemzésének és értékelésének továbbfejlesztése kapcsán. Ráműtöttem a határvédelem ellenőrző rendszereinek fejlesztési lehetőségeire az MI és a nagy nyelvi modellek által. Kiemeltem a személyazonosítás és azonosítást követő ellenőrzés fejlesztési metodikáit a jelenleg is működő MI-t használó rendszerek mentén, mint például a KIOSK. Az adat alapú komplex elemzéseket használó bűnüldözés és bűnmegelőzés kapcsán kifejtettem a rendészeti adatgyűjtés fontosságát azzal, hogy rámutattam a ChatGPT korlátaira, miszerint személyes adatok gyűjtésére a program algoritmusai alkalmatlanok. Alkalmos lehet ugyanakkor egyfajta rendészeti monitoring folyamatra, amely útján jogsértések előzhetőek meg, illetve kezelhetőek a nyelvi modell által. Az MI veszélyeire külön felhívom a figyelmet az MI hallucinációi kapcsán. Kiemeltem a chatbotok lakossági információszolgáltató képességét, amely nagyban hozzájárulna a rendőrségi alkalmazottak

tehermentesítéséhez, például az automatizált hívásfogadó rendszerek esetében. Kihangsúlyoztam a ChatGPT és a helyi kriminálstratégiai adatok gyűjtését és rendszerezését, mivel fontosnak látom a nagy nyelvi modell alkalmazhatóságát a lokális stratégiák tervezésében. Külön foglalkoztam a nagy nyelvi modell alkalmazásával összefüggő támogató funkciók fejlesztésével, igazgatási és jogalkotási tevékenységgel. Ennek során olyan további lehetséges funkciók bevezetésére tettem javaslatot, mint a programautomatizálás, a természetes nyelvfeldolgozás kiépítése vagy a hangvezérlés. A nagy nyelvi modellek jogszabályalkotási képességeinél külön kitértem arra, hogy a jelenlegi EU-s jogszabályok nem engedik még az MI e célú alkalmazását, ugyanakkor előkészítői folyamatokhoz való igénybevétele praktikus lehet. Az MI alapú bűnelkövetői kapcsolati hálózatok feltérképezése és a gráf adatbázis további fejlesztése kulcsfontosságú az eredményesebb bűnüldözés szempontjából. Az MI alapú elemzés lehetővé teszi a bűnözői kapcsolatok gyors és hatékony azonosítását, ami segíti a rendőrséget a bűnügyek hatékonyabb felderítésében és megelőzésében. Az MI alapú elemzés és az adatbázis további fejlesztése együttesen segítheti a rendőrséget a bűnözői kapcsolatok eredményesebb feltérképezésében. Az MI technológiák, mint a ChatGPT integrálása a nyomozási folyamatba lehetővé teszi a nagyobb hatékonyságot és az eljárási idő felgyorsítását. Mint korábban említettem, az MI technológiákat alkalmazhatják adatelemzésben, bűnözői kapcsolatok feltérképezésében és a rendőrök hatékonyabb támogatásában. Tanulmányomban külön foglalkozom a prediktív rendszettel: megállapítom, hogy az adatelemzés és a gépi tanulási algoritmusok kombinációjára épülő prediktív rendszer célja azon területek azonosítása, ahol a bűncselekmények valószínűsége a legmagasabb. Felhívom a figyelmet azonban arra, hogy az algoritmusokban rejlő elfogultság és megkülönböztetés lehetősége aggodalomra ad okot. Ennek oka, hogy a ChatGPT-t az emberek nagy mennyiségű adattal töltik fel, ezenkívül fennáll a hamis jelzések veszélye, valamint az is, hogy megfelelő szakértelem híján emiatt az MI bizonyos területeken túlzott rendőri jelenlétet javasolna. Az automatizált szövegleírással, a felolvasással és az iratkezeléssel, valamint az ügyfélkapcsolati rendszer fejlesztésével kapcsolatban is javaslatokat tettem. Az MI technológiák automatizálása és gépi tanulása hozzájárulhat a rendőrségi jelentések hatékonyabb és pontosabb előállításához. Az MI alapú eszközök segítségével a rendőrök sokkal gyorsabban és hatékonyabban írhatnak jelentéseket. Az MI technológiák segítséget nyújthatnak az információk automatikus adatbázisba vagy rendszerbe történő importálásában és elemzésében. Ugyanakkor fontos hangsúlyozni, hogy az MI

technológiák nem helyettesíthetik teljesen a rendőrök szerepét a jelentéskészítésben. Az automatizált szövegleíratkozás lehetővé teszi a beszélgetések egyszerűbb és gyorsabb kiértékelését. A ChatGPT kiválóan teljesíti a háttérzajok eltávolításában és több ember hangjának megkülönböztetésében. Az automatikus szövegfelolvasók hasznos eszközök lehetnek a rendőrség számára, amikor nyomozásokhoz szükséges információkat kell gyűjteni és értelmezni. Ezen szövegfelolvasók segítségével a rendőrség könnyen szűrheti az online platformokon zajló bűncselekményekkel, támadásokkal, fenyegetésekkel és illegális tevékenységekkel kapcsolatos beszélgetéseket. Az MI, mint például a ChatGPT alkalmazása számos módon hozzájárulhat a rendészeti okmányokkal kapcsolatos adatkezelés és feldolgozás támogatásához. Az MI képes lehet a digitális okmányok, mint például személyi igazolványok, útlevelek vagy hatósági jelzések automatizált kezelésére és ezek digitalizálására. A ChatGPT alkalmazása a rendészeti területen a kompetenciafejlesztés, a képzések és továbbképzések szempontjából is jelentős lehet. Az ilyen chatbotok széles körű rendőrségi információkat és szolgáltatásokat tárolnak, amelyek segíthetik a rendőröket készségeik fejlesztésében. Ezek a chatbotok hozzáférést biztosítanak olyan információkhoz és eszközökhöz, amelyek segítik a rendőröket a tények és trendek jobb megértésében, valamint hatékony kezelésében. Ezenkívül a chatbotok segíthetnek a rendőrök képzésében is, olyan készségek fejlesztésében, mint az interperszonális kommunikáció, a helyzetmegoldás, a jogi ismeretek és bűnüldözési stratégiák. Külön kiemelném a ChatGPT mint rendőrségi mentor lehetőségét. A ChatGPT integrálása a rendőrségi infrastruktúrába komoly felelősséget ró a rendőrségre a biztonságos és megbízható működésért. Az adatvédelemnek kiemelt figyelmet kell szentelni a megfelelő szabályok, biztonsági szabványok és eljárások kialakításával. Az infrastruktúrának képesnek kell lennie a chatbotok és virtuális asszisztensek karbantartására és fejlesztésére a rendőrség által kezelt adatokkal hatékonyan való működés érdekében. A jövőben várható, hogy az MI jelentős hatással lesz a rendvédelemre, segítve a hatékonyabb bűnözés elleni küzdelmet és a közrend fenntartását, ugyanakkor számos komplex kihívást is felvet. Az egyensúly megtalálása az MI alkalmazása és a lakosság jogai között kulcsfontosságú lesz a rendvédelmi területen.

Irodalomjegyzék

- Bagó, P. (2023) *Kiberbiztonság és a mesterséges intelligencia kapcsolata*. *Gazdaság és Pénzügy*, 10 (2). pp. 196-221. ISSN 2415-8909.

- Balla, J.–Kui, L. (2017) A határőrizeti célú ideiglenes biztonsági határzár és határőrizetre gyakorolt hatásai. *Hadtudományi Szemle*, 1. sz. pp. 222-238.
- Berthel, R.–Lapp, M. (2017) *Kriminalstrategie*. Heidelberg, C. F. Müller GmbH.
- Bialkó, L.G. (2023) [Mesterséges Intelligencia: Akkor most az MI tényleg elveszi a munkánkat?](#) *Népújság (Murska Sobota)*, Vol. 67.
- Bicskei, T. (2023) *A mesterséges intelligencia közigazgatásában való felhasználásával okozott kár*. *Közigazgatástudomány*, 3 (1). pp. 99-114. ISSN 2786-1910.
- Bogdány, Gy. (2018) Bűncselekmény-sorozatokat megszakítása, bűnözői csoportok bomlasztása. *Belügyi Szemle*, 66(7-8). pp. 168-180.
- Borbíró, A.–Kerezi, K. (szerk.) (2009) *A kriminálpolitika és a társadalmi bűnmegelőzés kézikönyve*. Budapest, Fresh Art Design Kft.
- Demeter, Z.–Mező, K. (2023) *Tanító szakos hallgatók és a mesterséges intelligencia = Lower-elementary school teacher professional students and artificial intelligence*. *Mesterséges Intelligencia*, 5 (1). pp. 73-87. ISSN 2676-9611.
- Deng, J.–Lin, Y. (2022) The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, 2(2). pp. 81-83.
- Déri, A. (2022) *Drónok alkalmazhatóságának lehetőségei a rendőrségen*. *Rendvédelem: A Belügyi Tudományos Tanács Online Folyóirata*, 11 (2). pp. 18-32.
- Dobó, J.–Gyaraki, R. (2021) *A mesterséges intelligencia egyes felhasználási lehetőségei a rendvédelmi területeken = Some Uses of Artificial Intelligence in Law Enforcement*. *Magyar Rendészet: A Nemzeti Közszolgálati Egyetem Rendészettudományi Szakmai Folyóirata*, 21 (4). pp. 67-81. ISSN 1586-2895.
- Farkas, K. (2023): *Mesterséges Intelligencián alapuló eljárások alkalmazása infokommunikációs hálózatokban*. *Híradástechnika (1962)*, 78: 1 pp. 32-39.
- Fejes, Zs.–Klotz, P.–Csóka, G. (2014) *Biztonság és bizalom a kormányzatban. A jó állam mérhetősége*. Budapest, Nemzeti Közszolgálati Egyetem.
- Fenyvesi, Cs. (2004) A XXI. századi bűnüldözés-tudomány nemzetközi tendenciái. *Magyar Tudomány*, 49: 6. pp. 757-764.
- Finszter, G. (2003) *A rendészet elmélete*. Budapest, KJK Kerszöv. Jogi és Üzleti Kiadó.
- Finszter, G. (2012) *A rendőrség joga*. Budapest, Duna Mix Kft.
- Finszter, G. (2016) *Bűnüldözés és Rendészeti modellek*. In: Borbíró, A.–Gönczöl, K.–Kerezi, K.–Lévai, M. (szerk.): *Kriminológia*. Budapest, Wolters Kluwer Kiadó.
- Finszter, G. (2018) *Rendészettan*. Budapest, Dialóg Campus Kiadó.
- Fórizs, S. (2014) *A stratégiaalkotás, a változásmenedzsment folyamata, elemei, szervezetfejlesztés a rendészeti szervezeteknél. Projektmenedzsment*. In: *A Rendészeti Szervek vezetése és szervezésmélete*. Budapest, Nemzeti Közszolgálati Egyetem, Rendészettudományi Kar. pp. 11-29.
- Gyaraki, R. (2023) *A mesterséges intelligencia felhasználási lehetősége és fejlesztésének szükségessége a jogalkalmazásban*. In: Kovács, Z. (szerk.) *A mesterséges intelligencia és egyéb felforgató technológiák hatásainak átfogó vizsgálata*. Budapest, Katonai Nemzetbiztonsági Szolgálat, pp. 393-422. ISBN 978-615-6128-18-8.
- Gulyás, A. (2023). *A nyílt forrásból származó adatgyűjtés automatizálásának lehetőségei*. *Belügyi Szemle*, 71(7), pp. 1237-1269.
- Harries, K. (1999) *Mapping Crime. Principle and Practice*. Washington, CMRC, National Institute of Justice.
- Herke, Cs. (2021) *A mesterséges intelligencia kriminalisztikai aspektusai*. *Belügyi Szemle*, 69(10), pp. 1709-1724.
- Herke, Cs. (2023) *Mesterséges intelligencia a büntetőjogi döntéshozatalban*. *Jogtudományi Közlöny*, 2023/4. pp. 165-176.
- Hertelendi L.–Hornyik, Z. (2022) *Mesterséges intelligencia a köz szolgálatában: Interjú Hajzer Károly informatikai helyettes államtitkárral*. *Belügyi Szemle*, 70(1). pp. 207-217.
- Iványi, M.P. (2023) *Az infokommunikációs technológiák és a mesterséges intelligencia intézményeinek nemzetközi politikai gazdaságtani keresztszete*. *Közigazgatástudomány*, 3 (1). pp. 51-67. ISSN 2786-1910.
- Kalmár, Á (2018) *Innovációs javaslatok a határrendészeti szolgálati ág részére a tömeges méretű migráció kezelésében*. *Rendőrségi Tanulmányok* 2018/1. sz. pp. 75-100.
- Kis, N.–Szenes, Z.–Vajda, A.–Tálas, P.–Schweickhardt, G.–Fejes, Zs.–Klotz, P.–Csóka, G. (2014) *Biztonság és bizalom a kormányzatban. A jó állam mérhetősége*. Budapest, Nemzeti Közszolgálati Egyetem.
- Kis, N. (2016) *Biztonság, Bizalom és a jó kormányzás. Államtudományi Műhelytanulmányok*, 2016. évi 23. sz. pp. 2-13.
- Kiss, Cs. (2023) *Világuralmi verseny az elsőbbségért a mesterséges intelligenciában*. *Hadtudományi Szemle*, 16 (2). pp. 211-220. ISSN 2676-9816.
- Kollár, Cs. (2018) *A mesterséges intelligencia kapcsolata a humán biztonsággal*. *Nemzetbiztonsági Szemle*, 6. évf. 1. sz. pp. 5-23.
- Kuhn, J. (2023) *Data Ethics and the Dilemma Created by Turing's Learning Machines*. University of Northern Iowa.
- Mátyás, Sz.–Sallai, J. (2014) *Kriminálgeográfia*. In: Prof. Dr. Ruzsonyi P. (szerk.): *Tendenciák és alapvetések a*

- bűnügyi tudományok köréből. Budapest, Nemzeti Közszerkesztési és Tankönyv Kiadó. pp. 335-356.
- Mazsu, D. (2019) Jog és a mesterséges intelligencia: Új szereplő, régi alakzatok? (Gondolatok Jacob Turner könyve kapcsán). *Pro Futuro*, 9(1), pp. 137-145.
- Mezei, K. (2023) [A mesterséges intelligencia jogi szabályozásának aktuális kérdései az Európai Unióban](#). In *Medias Res: Folyóirat A Sajtószabadságról És A Médiaszabályozásról*, 70: 1 pp. 53-70.
- Necz, D. (2020) *A mesterséges intelligencia belügyi és biztonsági célú alkalmazása*. *Scientia et Securitas*, 1 (1). pp. 49-53. ISSN 2732-2688.
- Necz, D. (2022) *A mesterséges intelligencia felhasználásával történő adatkezelések egyes sajátos szempontjai*. *Acta Humana: Hungarian Centre For Human Rights Publications*, 10 (3). pp. 95-123. ISSN 0866-6628.
- Négyesi, I. (2023) *A mesterséges intelligencia társadalmi és etikai kérdései*. *Honvédségi Szemle: A Magyar Honvédség Szakmai, Tudományos Folyóirata*, 151 (4). pp. 6-18. ISSN 2060-1506.
- Németh, A.–Virágh, K. (2023) *Mesterséges intelligencia és haderő - Katonai alkalmazási lehetőségek VII. rész*. *Haditechnika*, 57 (1). pp. 2-6. ISSN 0230-6891.
- Pintér, M. (2020) *A mesterséges intelligencia etikai aspektusai*. *Rendvédelem: a Belügyi Tudományos Tanács online folyóirata*, 2020/1. lapszám. pp. 27-36.
- Ritecz, Gy.–Sallai, J. (2016) *A migráció trendjei, okai és kezelésének lehetőségei 2.0*. Budaörs, Hanns Seidel Alapítvány.
- Sallam, M. (2023) ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare*, Vol. 11. No. 6. p. 887. MDPI.
- Segal, M. (2018). Automatic pilots – more robotics and artificial intelligence in the workplace doesn't have to destroy your job. *Nature*, 563(Nov). pp. 132-135.
- Szabó, I. (2019) Automatizált döntéshozatal és a büntetőeljárás. *Ügyészek Lapja*, 2019. 4-5. Sz. pp. 5-20.
- Sütő, É. (2023) *A mesterséges intelligencia geopolitikája*. *Külügyi Szemle*, 22 (2). pp. 144-150. ISSN 1587-9089.
- <https://forbes.hu/interju/mesterseges-intelligencia-chatgpt-interju/> [Letöltve: 2023.09.14.].
- Nagy, N. (2023) [A Rubiconon túl: a mesterséges intelligencia emberi jogi kockázatai](#). *Ludovika Blog*. <https://www.ludovika.hu/blogok/kormblog/2023/02/03/a-rubiconon-tul-a-mesterseges-intelligencia-emberi-jogi-kockazatai/> [Letöltve: 2023.09.20.].
- Szabó, I. (2011) *A rendészettudomány múltja, jelene és jövője*. <https://docplayer.hu/46313059-A-rendeszettudomany-multja-jelene-es-jovoje.html> [Letöltve: 2023.09.27.].
- URL1: <https://www.globaltimes.cn/page/202307/1293835.shtml> [Letöltve: 2023.09.09.].
- URL2: <https://www.europol.europa.eu/publications-events/publications/chatgpt-impact-of-large-language-models-law-enforcement#downloads> [Letöltve: 2023.09.11.].
- URL3: *Europol: ChatGPT. A nagy nyelvi modellek hatása a bűnüldözésre*. <https://www.europol.europa.eu/publications-events/publications/chatgpt-impact-of-large-language-models-law-enforcement#downloads> [Letöltve: 2023.09.11.].
- URL4: *Magyarország Mesterséges Intelligencia Stratégiája 2020-2030* <https://digitalisjoletprogram.hu/files/2f/32/2f32f239878a4559b6541e46277d6e88.pdf> [Letöltve: 2023.09.07.].
- URL5: <https://openai.com/blog/introducing-chatgpt-and-whisper-apis> [Letöltve: 2023.09.19.].
- URL6: <https://platform.openai.com/docs/models/continuous-model-upgrades> [Letöltve: 2023.09.19.].
- URL7: <https://www.digitalhungary.hu/e-volution/A-jogalkotas-csak-lohol-a-technologia-utan-mi-var-Europaban-a-ChatGPT-re/20347/> [Letöltve: 2023.09.17.].

Hivatkozott jogszabályok, normák, vitaanyagok

Internetes hivatkozások

- Dunavölgyi, Sz. (2011) *Rendészet – rendvédelem a magyar jogalkotásban („A rendészettel kapcsolatos rendszerelméleti fogalmak tisztázása” c. kutatás része)*. <https://bm-tt.hu/wp-content/uploads/2022/02/rendeszett-vs-rendvedelem-dunavolgyi.pdf> [Letöltve: 2023. 09.09.].
- Kis, J. (2023) *Hallucináló csetbotok vagy mindentudó AI? Interjú a chatgpt gyerekbetegségeiről*.

- 115/2003. OGY határozat a bűnmegelőzés nemzeti stratégiájáról.
- A Magyar Köztársaság Rendészeti Stratégiája. 2009. január 15. munkadokumentum
- A Robotzsaru rendszer alkalmazás-támogatásáról, fejlesztéséről és a használatához szükséges képzésről szóló 13/2023. (V.31.) ORFK utasítás.*
- Az informatikai alkalmazásfejlesztési folyamatok szabályozásáról szóló 14/2022. (IV.7.) ORFK utasítás.*
- A Rendőrségről szóló 1994. évi XXXIV. törvény.
- Igazságügyi és Rendészeti Minisztérium: *Rendészeti Stratégia*. Készült a 2006. szeptember-októberi fővárosi demonstrációkkal, utcai rendezvényekkel és

rendfenntartó intézkedésekkel kapcsolatos
eseményekről készült vizsgálati jelentés alapján

végrehajtandó feladatokról szóló 1013/2007. (III. 13.)
Korm. határozat 4. pontja alapján.

Kérdezz bármit, tudom a választ! – avagy a ChatGPT adatvédelmi kérdései, különös tekintettel a személyes adatok védelmére és a kockázati besorolásra

Ask me anything, I know the answer! – ChatGPT's data protection issues, with a special regard to personal data protection and risk classification

DOI: [HTTPS:// DOI.ORG/10.53793/RV.2024.1.4](https://doi.org/10.53793/RV.2024.1.4)

Absztrakt

Jelen cikkünk témája a ChatGPT, mint nagy nyelvi modell adatvédelmet és kockázati besorolást érintő vizsgálata. Ezen vizsgálódás elengedhetetlen kellékeként először magát a ChatGPT alapjául szolgáló technológiát tekintjük át dióhéjban, beleértve a nagy nyelvi modell működését biztosító és/vagy segítő támogatószolgáltatásokat. Ezen összetett rendszert pedig a GDPR-szabályai segítségével vizsgáljuk, az egyes lényeges adatvédelmi kérdésekre kitérve. Mindezek alapján pedig megkíséreljük behatárolni a kockázati szintet, ami a ChatGPT, mint rendszer veszélyességét jelenti a felhasználókra nézve. Ezután valós jogeseteken keresztül nézzük meg, éles helyzetben, hogy jelen állás szerint milyen megoldásokra lehet számítani a mesterséges intelligencia jogsértő magatartásával szemben. A cikk lényege tehát, hogy szélesebb képet adjon a nagy nyelvi modellek jelenlegi adatvédelmi veszélyeiről, kockázati szintjéről, és azokról a jogi utakról, megoldásokról, amelyek jelen helyzetben a jogalkalmazók kezében vannak a mesterséges intelligencia rendszerekkel szemben. A cikk célja, hogy röviden összefoglalja hol is állunk jelenleg, és egyfajta javaslatot tegyen, merre érdemes tovább haladni.

KULCSSZAVAK: CHATGPT, MESTERSÉGES INTELLIGENCIA, NAGY NYELVI MODELLEK, ADATVÉDELEM, SZEMÉLYES ADAT

Abstract

The topic of this article is the examination of the data protection issues of ChatGPT as a large language model, and its risk classification. Firstly, we are considering an essential accessory to this study that is the underlying technology of ChatGPT itself in its entirety, including the support services that ensure and/or support the functioning of the large language model. Secondly, we are going to look at this complex system through the rules of the GDPR and each of the key data protection issues. Based on all of this, we are trying to narrow down the level of risk chat ChatGPT poses as a system to its users and then we are examining what kind of solutions can we expect from the current state of affairs to deal with the illegal behaviour of artificial intelligence. The purpose of this article is therefore to provide a broader picture of the current data protection threats of the big four models, their level of risk, and the legal avenues and solutions available to law enforcement agencies in the current situation in relation to AI systems.

KEYWORDS: CHATGPT, ARTIFICIAL INTELLIGENCE, LARGE LANGUAGE MODELS, DATA PROTECTION, PERSONAL DATA

Bevezetés – A ChatGPT és az adatvédelem

A mai világban is megállja a helyét a mondás, miszerint a tudás hatalom, ezt ChatGPT működése alapján átalakíthatjuk akár arra, hogy az „információ hatalom” vagy „az adat hatalom”. A ChatGPT, mint nagy

nyelvi modell tanításához felhasznált adatok fellelhetősége, ellenőrzése és azok felhasználása az átlag felhasználó számára nem transzparens. A chatbot használata oly mértékben könnyíti meg a mindennapi

munkát, hogy lassan a már jól megszokott keresőprogramokkal vetekszik népszerűsége. Az előbb említett egyszerűség és népszerűség az egyik legnagyobb veszély akkor, amikor a színpalak mögé nézünk és elolvassuk az „apróbetűs részeket”. Felmerülnek olyan kérdések, hogy: „Kinek az adataival dolgozik?” „Meddig tárolja a megszerzett adatokat?” „Az illető hozzájárul-e az adatok megszerzéséhez?” és még sorolhatnánk. Cikkünkben ezen kérdéseket boncolgatjuk. Nem a mindenki számára egységesen elfogadható válasz megtalálásán van a hangsúly, sokkal inkább az egyes problémák bemutatásán, azok technikai és jogi megközelítésén. A cikk egy technológiai és egy jogi adatvédelmi részből áll. A technológiai áttekintést a mesterséges intelligencia meghatározásával indítjuk, majd pedig ebből kiindulva lépésről lépésre eljutunk a ChatGPT-ig. Ezen lépések során érinteni kívánjuk többek között a gépi tanulást, a mélytanulást, a neurális hálót, a modellt és a különböző tanítási technikákat. A végére pedig maga a GPT, illetve a ChatGPT lehetséges problémái, hibái kerülnek tárgyalásra. Ezt követően a GDPR és a ChatGPT működése során felmerülő jogi kollíziót vesszük górcső alá, kiegészítve az egyes tagállamok ChatGPT által kiváltott reakcióival.

Mi a ChatGPT?

Mesterséges intelligencia és gépi tanulás – mi különbözteti meg őket?

A mesterséges intelligencia fogalmának meghatározását az AI Act módosított változatában használt definícióval határozzuk meg, ami nem más, mint: „(AI rendszer): olyan gépi alapú rendszer, amelyet úgy terveztek, hogy különböző szintű autonómiával működjön, és amely explicit vagy implicit előrejelzésekre, ajánlásokra vagy döntésekre képes, amelyek befolyásolják a fizikai vagy virtuális környezetet” (AI Act). A mesterséges intelligencia célját ugyanakkor több vezető mesterséges intelligencia kutató cég szerint az emberi intelligencia kognitív képességekkel bíró gépek létrehozására való törekvése jellemzi. Ezek a képességek az IBM szerint: az arc- és beszéd felismerés, a döntéshozatal, a fordítás; a Google idesorolja még ezen kívül a látást, az adatelemzést, az ajánlások adását. Ezek mind egy-egy területét vagy szeletét adják a mesterséges intelligenciának, mint interdiszciplináris területnek. Ahhoz, hogy ezekkel a képességekkel egy mesterséges intelligencia rendelkezzen, többféle megközelítésre van szükség, ezért a mesterséges intelligencián belül 6 klasszikus területet különböztetünk meg:

1. Gépi tanulás
2. Gépi látás
3. Neurális hálózatok

4. Mélytanulás
5. Természetes nyelvi feldolgozás
6. Kognitív számítástechnika.

A gépi tanulás a mesterséges intelligencia olyan részhalmaza, amely automatikusan képessé tesz egy gépet vagy rendszert arra, hogy tapasztalataiból tanuljon és fejlődni tudjon. A nagymennyiségű adat elemzésére a gépi tanulás explicit programozás helyett algoritmusokat használ a felismerésekből való tanulásra, valamint a döntések meghozatalára. A gépi tanulási algoritmusok idővel javítják teljesítményüket, ahogyan egyre több adatot használnak fel. Ha az emberi kognitív képességekre szeretnénk vele reflektálni, akkor a gépi tanulás egyértelműen megoldással szolgál az adatelemzéshez, ajánlások megtételéhez.

A mesterséges intelligencia alapvető célja, hogy az ember kognitív képességeinek teljes repertoárját a gépi működés által képezze le. Ehhez képest a gépi tanulás egy bizonyos területnek csak egy-egy szeletét képes lehetővé tenni. Ezért a mesterséges intelligencia tágabb fogalom, amelyen belül a gépi tanulás elhelyezkedik.

A gépi tanulás, neurális hálózatok, mélytanulás elhatárolása

A mesterséges intelligencia működésén belül többféle megoldást találunk, amely lehetővé teszi a rendszer működését. Ezeket úgy kell elképzelni, mint egymásba ágyazott megoldásokat, elveket. Az egyik ilyen megoldás a neurális háló. Ez a név akár a biológiából is ismerős lehet, hiszen az emberi agyat, az idegrendszert alapul véve végzi el a szükséges számításokat. Természetesen ezekből a hálók közül többféle létezik, amelyeket többek között a felépítés, a mélység, az információ iránya és egyéb tényezők alapján különböztetünk meg. Más neurális hálózat, más megoldást kínál bizonyos problémákra, ezek a megoldások a gépi tanulás és a mélytanulás. A gépi tanuló algoritmusok családján belül helyezkedik el a mélytanulás is, amely egy olyan válfaja a gépi tanulásnak, ahol az adott mélytanuló mesterséges intelligencia önmagától, külső segítség nélkül fedezi fel a tanulóadatokban lévő különböző mintákat, szabályosságokat. Ez csak bonyolultabb felépítésű, több rétegű neurális hálóval lehetséges, mint a gépi tanulás esetén, ahol a gépi tanuló rendszer címkézés segítségével tanul.

Architektúrák és modellek – mik ezek?

A mélytanulás – ahogy fent írtuk – neurális hálóból épül fel, ezen neurális hálók építőkövei a következők:

1. bemenet

2. súly
3. átviteli függvény
4. aktiválási függvény
5. torzítás.

Ezekből az építőkövekből többféle felépítésű neurális hálózat (architektúra) hozható létre attól függően, hogy milyen problémát szeretnénk megoldani, milyen felhasználásban gondolkodunk. Léteznek úgynevezett szabványos neurális hálózat architektúrák, avagy neurális hálózat felépítési tervek.

A neurális hálózat architektúráktól részben el kell különíteni a modelleket, hiszen a gépi tanuló modell, már a tanult adatok alapján képes előrejelzést készíteni, döntést hozni új, korábban nem látott adatkészletből. A gépi tanuló modell megalkotása magában foglalja egy neurális hálózat architektúra elkészítését, majd egy adatkészlet alapján történő betanítását, addig a pontig, ahonnan már önállóan és megbízhatóan el tudja végezni feladatát.

Ezek után rátérhetünk a foundation model, avagy az *alapmodell* kérdésére. Az AI Act módosított változata kiter az alapmodell meghatározására is: „Az alapmodellek új keletű fejlődés, amelyben a mesterséges intelligenciamodelleket olyan algoritmusokból fejlesztik ki, amelyek célja a kimenet általános és sokoldalúságának optimalizálása. Ezeket a modelleket gyakran adatforrások széles skáláján és nagy mennyiségű adaton képzik ki, hogy a downstream feladatok széles skáláját végezzék el, beleértve olyanokat is, amelyekre nem kifejezetten fejlesztették és képezték ki őket...[...].” (AI Act. (6oe)). Ez alapján elmondhatjuk, hogy az alapmodell olyan neurális hálózat, amelyet általában nyers adatokon képeznek ki, felügyelet nélküli tanulással, és többféle feladat elvégzésére adaptálható, olyanokra is, amelyekre célzottan nem fejlesztették ki őket.

A *transformer modell* egy neurális hálózati architektúra, amely képes automatikusan átalakítani egy adott típusú bemenetet egy másik típusú kimenetű. A kifejezést egy 2017-es Google-dokumentum alkotta meg, amely megtalálta a módját egy neurális hálózat betanításának az angolról franciára való fordításra, és a többi neurális hálózat képzési idejének egynegyedével. A transformer működését néhány mondatban a következőképpen foglalhatjuk össze: a transformer alapfelépítése szerint kódoló, és dekódoló részre bontható. Két alapvető fázisa van a modell tanításának. Az első fázisban egy transformer nagy mennyiségű címkézetlen adatot dolgoz fel, hogy megtanulja a nyelv szerkezetét vagy egy jelenséget. A modell betanítása után hasznos lehet finomhangolni egy adott feladathoz, ez a második fázis.

Az modellek tanítási technikái – hogyan jön létre egy modell?

Láttuk, hogy a neurális hálózat architektúráját meg kell alkotni és be kell tanítani, hogy így már alapmodellként kezelhessük tovább.

Előtanítás

Az előtanítás célja, hogy nagy mennyiségű adat alapján a modell megtanulja megjósolni a következő szót (nagy nyelvi modellek esetében). Ebben a fázisban a modell még nem tanul speciális feladatot, csupán a következő szó megjóslásának módját tanulja meg, egy adott szöveggörnyezetben. Ez a folyamat többféle tanítási mechanizmus szerint történhet, azonban a GPT esetében az autoregresszív tanítási módot használják, vagyis a modellnek úgy kell kitalálnia a következő szót, hogy egy adott mondatban nem látja a szöveg folytatását, így az előző szavakból kell kiindulnia. Például a ChatGPT-nek leírunk egy fél mondatot, majd azt befejezi: „A macska fel van” – a ChatGPT befejezte: „ A macska fel van készülve a kalandra”. Ezzel a fázissal létre lehet hozni a neurális hálóból az alapmodellt, tehát a transformer alapú neurális hálóból itt kapjuk meg a GPT-t, hiszen az előtanítási fázis elegendő, hogy generatív is legyen, de még nem lehet vele „csevegni”.

Finomhangolás

Az előtanítás után a kezünkben van egy úgynevezett PLM (Pre-trained language model), egy általános nyelvi tudással rendelkező modell. Ha konkrét feladatokra szeretnénk használni, akkor finomhangolnunk kell, amihez egyébként már jóval kevesebb adatra van szükség, mint az előtanításhoz. Ennek során egyrészt plusz feladatspecifikus rétegeket adhatnak a modell neurális hálózatához, illetve egy kisebb adatkészlettel betanítják kimondottan egy konkrét feladatra. A GPT így tanult meg kvázi chatbotként beszélgetni, mivel alapjában véve a GPT csak egy szöveggeneráló, szöveg értelmező és fordító modell.

Generatív modell

A generatív modellek lényege, hogy megértsék, rögzítsék a mögöttes minták eloszlását egy adathalmazból, ezután a modell új adatokat generálhat, amelyek hasonló tulajdonsággal rendelkeznek, mint az eredeti adatok. Fontos, hogy a generatív modell egy alapmodell finomhangolása után nyeri el végső „formáját”, tehát azt a feladatot, amit a

modellnek szánunk, csak a finomhangolás után képes megfelelően ellátni.

ChatGPT

Mindezek fényében leránthatjuk a leplet a ChatGPT-ről, amely nem más, mint egy generatív módon működő, előtanított/általános tudással ellátott, transformer modell, chatbotként finomhangolva és működtetve.

A transformer modell kockázatai

A transformer modelleknek is megvannak a maguk kockázatai, amivel számolni kell a modell teljes életciklusában, kezdve a fejlesztéstől, egészen a használatig. Három példát említünk, amikről úgy gondoljuk fajsúlyosak a kockázatok között.

Felmerülhet az úgynevezett „adatmérgezés” esete, amikor a támadó rosszindulatú példákat helyez el a modell tanítási adatbázisában, amivel hatással tud lenni a modell minden alkalmazására, oly módon, hogy ezzel a modell teljesítményét rombolja.

Ha az alapmodell például egyetlen vállalat *privát adataival van előtanítva*, akkor fennáll a veszélye, hogy az összes downstream⁷ alkalmazásra ezen adatokat felfedje. Például, ha valamit kérdezzünk a rendszertől, akkor van rá esély, hogy a modell válasza bizalmas információkat fog tartalmazni az adott vállalatról.

Problémát jelenthet még az úgynevezett *kettős felhasználás*. Ezt a problémát az alapmodellek rugalmassága adja, vagyis ezek a modellek több feladathoz is viszonylag könnyen alkalmazkodnak. Ez viszont lehetővé teszi, hogy eredeti céljuktól eltérően használják őket. Például egy modell, amely eredetileg kép-szöveg párok előrejelzésével foglalkozott, egy idő után képes volt nagy részletességgel felismerni az arcvonásokat. Az alapmodellek könnyű adaptálhatósága tehát lehetőséget teremthet akár a visszaélészerű használatra is.

A ChatGPT adatvédelmi vonatkozásairól, figyelembevéve az egyéb ún. támogatószolgáltatásokat

Az előzőleg tárgyalt technikai részben a GPT modell-re helyeztük a hangsúlyt. A modellek neurális háló alapján működnek, amelyeknek nincs a klasszikus értelemben vett memóriájuk. Így máshol kell keresni azt a rendszert, ahol az adatokat tárolják és feldolgozzák, éppen ezért ebben a fejezetben először a ChatGPT működését elősegítő egyéb rendszerekről lesz szó első körben, majd rátérünk jogi értékelésére.

Az OpenAI honlapján öt olyan szereplőt találhatunk (URL), akiket a cég úgynevezett alfeldolgozókként tüntet fel, azonban ezen öt vállalat tevékenysége nagyon eltér, így adatvédelmi kérdésekben sem lesznek egységesen relevánsak:

- Cloudflare: ezen vállalat az úgynevezett CDN-szolgáltatással segíti az OpenAI-t, vagyis segít a tartalomszolgáltatásában
- Microsoft: a Microsoft az Azure felhőszolgáltatással csatlakozik be az OpenAI szolgáltatási körébe
- OpenAI affiliates: szolgáltatások és támogatások
- Snowflake: a Data Lake és a Data Warehousing szolgáltatásokkal támogatja az OpenAI szolgáltatásait
- TaskUS: felhasználói támogatás és felügyelet.

Egyéb közreműködő szolgáltatások

- Cloudflare: mint említettük, az OpenAI a Cloudflare egy konkrét szolgáltatását, a CDN-t nevezi meg. A CDN nem más, mint egy szerverhálózat, amelynek célja, hogy a lehető leggyorsabban, megbízhatóbban és biztonságosabban szállítsa a tartalmat az eredeti szervertől a felhasználókhöz. Előnye, hogy az oldalak betöltési idejének gyorsaságát ez biztosítja, ezáltal a világ különböző pontjain sem kell perceket várni a weboldalra míg betölt. Működése lényegében abból áll, hogy bár a weblap tartalmát eredetileg a szolgáltató tárolja, a weboldal tartalmait mégis több szerveren gyorsítótárazzák, vagyis lemásolják és különböző szervereken tárolják. Ezek a szerverek tartalmazhatnak személyes adatot is.
- Microsoft Azure: a Microsoft Azure szolgáltatásain keresztül is elérhető az OpenAI több modellje, többek között a GPT 4, GPT – 35 – Turbo, DALL-E. Ezen modellekhez az Azure úgynevezett REST API-val biztosít hozzáférést. Elmondhatjuk, hogy REST API-n keresztül nem tárolódik információ az Azure-szerverein, viszont az OpenAI mégis alfeldolgozóként kezeli, vagyis más módon juthat adat, információ a Microsoft birtokába. Jó példa erre, ha egy felhasználó finomhangolni szeretne egy modellt, akkor a saját adatbázisát az Azure rendszere fogja tárolni. Ebben az adatbázisban nem szükségeszerű személyes adat szerepeltetése, de mivel ennek semmi akadálya nincs, így számolni kell vele.

⁷ feladat specifikus alkalmazás, egy chatbot esetében „beszélgetés”

- *Snowflake*: ehhez először két fogalmat kell tisztáznunk, az első a Data Lake (magyarosan „adattó”), második a Data Warehouse (magyarosan „adattárház”). Data Lake alatt olyan adattárolót értünk, amely nagymennyiségű nyers adatot tárol, annak eredeti formájában. Itt strukturált, nem strukturált és félig strukturált adatok is vegyesen előfordulhatnak, képi, szöveges és hang adatok is tárolhatók ezen rendszeren belül. Azon vállalatok számára előnyös, amelyek bár sok adatot gyűjtenek, de ezeket nem kell azonnal feldolgozniuk. Data Warehouse esetén már strukturáltabb adatbázisban dolgozhatják fel és kezelhetik adataikat a vállalatok. Az OpenAI a Snowflake szolgáltatását használja, vagyis adatait a Snowflake szerverein tárolja. Minden bizonnyal kombinálják a Data Lake és a Data Warehouse előnyeit, vagyis a nyers, feldolgozatlan adatokat a Data Lake-ben tárolják, majd ezeket egy Data Warehouseban feldolgozzák és előkészítik az elemzéshez vagy a további felhasználáshoz.
- *OpenAI affilates*: jelenleg nincs tudomásunk arról, hogy az OpenAI hirdetett volna meg efféle programot, így ennek most nincs relevanciája, ezért az adatvédelmi vizsgálatból kihagyhatónak ítéljük.
- *TaskUS*: a TaskUS esetében felmerült a kérdés: „mire gondolhatt a költő?”, mivel az OpenAI felhasználói támogatást említi, biztonsággal és felügyelettel. A vállalat oldalát böngészve olyan szolgáltatásra bukkantunk, név szerint: Digital Customer Experience, amelyben a TaskUS az ügyfél-vállalat ügyfeleivel kommunikál, lényegében ügyfélszolgálatként működik. Az oldalon azt írják, hogy hangalapon, e-mail-en, chaten és technikai támogatással lépnek kapcsolatba az ügyfelek ügyfeleivel. Illetve van egy másik szolgáltatás Securing your Content elnevezéssel, ahol többek között felhasználó által generált tartalmak moderálását lehet igénybe venni, de vállalják még a generatív AI által kreált tartalmak moderálását is.

Adatvédelem (GDPR)

Elkezdjük a vizsgálódást az adatvédelem, pontosabban *Az Európai Parlament és a Tanács 2016. április 27-i (EU) 2016/679 Rendelete (URL2)* a természetes személyeknek a személyes adatok kezelése tekintetében történő védelméről és az ilyen adatok szabad áramlásáról, a továbbiakban: GDPR szemüvegén keresztül.

Személyes adat

A GDPR 4. cikkének (1) értelmében személyes adatnak csak természetes személlyel összefüggő adatok minősülnek, a szervezetek adatai nem tartoznak a szabályozás alá. Fontosnak tartjuk kiemelni, hogy adatvédelmi jogi szempontból az egyéni vállalkozó is természetes személynek minősül, tehát az egyéni vállalkozóval kapcsolatba hozható adat személyes adat. A személyes adat GDPR szerinti fogalma így hangzik: „azonosított vagy azonosítható természetes személyre („érintett”) vonatkozó bármely információ; azonosítható az a természetes személy, aki közvetlen vagy közvetett módon, különösen valamely azonosító, például név, szám, helymeghatározó adat, online azonosító vagy a természetes személy testi, fiziológiai, genetikai, szellemi, gazdasági, kulturális vagy szociális azonosságára vonatkozó egy vagy több tényező alapján azonosítható”.

Ha olyan módon kezeljük az adatokat, amelyek megakadályozzák, hogy ezen adatok a természetes személlyel kapcsolatba hozhatók legyenek, úgy megvalósul a GDPR 4. cikkének (5) bekezdése: „álnevesítés”: a személyes adatok olyan módon történő kezelése, amelynek következtében további információk felhasználása nélkül többé már nem állapítható meg, hogy a személyes adat mely konkrét természetes személyre vonatkozik, feltéve, hogy az ilyen további információt külön tárolják, és technikai és szervezési intézkedések megtételével biztosított, hogy azonosított vagy azonosítható természetes személyekhez ezt a személyes adatot nem lehet kapcsolni.

Mindezek alapján az OpenAI esetét megkíséreljük értelmezni, vagyis tárol-e személyes adatot, és melyek azok a vállalatok, amelyek az OpenAI által begyűjtött személyes adatokat tárolják. Fontos azzal kezdeni, hogy az OpenAI honnan és miként gyűjti be a személyes adatokat. Ilyen adatokként azonosíthatók:

- a) felhasználói fiók információk
- b) felhasználói tartalom
- c) kommunikációs információk
- d) közösségi média információk
- e) tanulmányok.

Az OpenAI szerverei nyilvánvalóan megkapják ezen adatokat, amennyiben a szolgáltatást használjuk, akár webes felületen, akár applikáción keresztül, hiszen, ha csak a felületet megnézzük, egyfajta történeti sávot találunk, ahol a rendszerrel folytatott, több korábbi kommunikációnk található meg. Így pedig az OpenAI képes a tartalmat személyhez kötni, ezzel személyes adatkezelés folyik.

Úgy gondoljuk logikus lenne, ha ezek a Snowflake szervereire kerülnének, hiszen az OpenAI részére biztosít Data Lake és Data Warehouse szolgáltatást egyaránt. Amint láttuk, a Data Lake-ben lényegében bármilyen adat, bármilyen strukturálatlanul is

eltárolható, így mivel kvázi nincs kezelve, a személyes adat jellegét nem veszíti el, bár ez a szolgáltatás akár több szerveren is megvalósulhat, mégis egyetlen összekapcsolt rendszerként létezik. A feloldást talán a Data Warehouse adhatja, ahol már az adatok egy begyűjtési és tisztítási, feldolgozási szakaszon mennek keresztül, így talán elérhetik az álnevesítési szintet – mivel külön szerveren, szűrve, más rendszerektől elkülönülten léteznek –, ha az érintettel csak további információk birtokában kapcsolható össze. Amennyiben, ezeket a már szűrt, tisztított és külön szerveren tárolt adatokat előtanításhoz használják fel, kérdés, hogy a modell rengeteg adat alapján mennyire képes összekapcsolni a természetes személyekkel, hiszen mint látjuk, a transformer modellek rendkívül jól képesek a szöveg-kontextusokat azonosítani.

A következő a Cloudflare esete, ahol a CDN szolgáltatás kap központi szerepet. A Cloudflare szerverei lépnek elsősorban kapcsolatba a felhasználó eszközével, ezáltal biztos, hogy IP-címeket és egyéb azonosítókat tároljanak a felhasználó eszközéről. Mindemellert azt a felületet is eltávolítja, amit a felhasználó előzőleg megtekintett, annak érdekében, hogy legközelebb gyorsabb adatátvitelt tegyen lehetővé. Mondhatjuk, hogy a Cloudflare szintén személyes adatot tárol.

A Microsoft esete egyszerűbb, hiszen a bemenetet nem tárolja a rendszer, mivel csupán REST API hozzáférést biztosít a modellekhez, azonban a felhasználói fiók adatait, illetve a felhasználó által feltöltött adatokat egy az egyben tárolják az Azure szerverei.

A TaskUS esetében e-mail cím, név, egyéb adatok azok, amik a megkeresésben (akár e-mail, akár chat vagy más) szerepelnek, minimum ennyi adatot tárol és kezel.

Álnevesítésről óvatosan beszéljünk ezen támogatószolgáltatások esetén, mivel bár biztosak lehetünk benne, hogy kódolják, titkosítják az adatokat, de egyazon adat egyszerre több helyen is előfordulhat, például az IP-cím, eszköz, név bemenet akár az OpenAI, akár a Cloudflare, akár a Microsoft, de talán még a Snowflake szerverein is megtalálható. Ezért az álnevesítés tökéletes megvalósításához erős kételyek fűződnek.

Az adatkezelés jogalapja

A jogalap kérdésében a GDPR 6. cikke rendelkezik, melyekből – véleményünk szerint – az OpenAI adatvédelmi szabályzata alapján leginkább az első teljesül a vállalat részéről: a) „az érintett hozzájárulását adta személyes adatainak egy vagy több konkrét célból történő kezeléséhez”.

Az érintett hozzájárulásnak GDPR szerinti definíciója: „az érintett akarátának önkéntes, konkrét és

megfelelő tájékoztatáson alapuló és egyértelmű kinyilvánítása, amellyel az érintett nyilatkozat vagy a megerősítést félreérthetetlenül kifejező cselekedet útján jelzi, hogy beleegyezését adja az őt érintő személyes adatok kezeléséhez” Ebből a definícióból néhány fontos és lényeges pont emelhető ki, amelyek a hozzájárulás kellékeit adják: önkéntesség, konkrét mivolt, megfelelő tájékoztatáson alapuló jelleg, és egyértelműség. Maga a hozzájárulás is meghatározott fogalom, amely 95/46/EK irányelv 2. cikkének h) pontja szerint: „az érintett kívánságának kinyilvánítása, amellyel beleegyezését adja az őt érintő személyes adatok feldolgozásához”.

Adatkezelő és adatfeldolgozó

Az adatkezelőt a GDPR 4. cikkének 7. pontja a következőképpen határozza meg: „az a természetes vagy jogi személy, közhatalmi szerv, ügynökség vagy bármely egyéb szerv, amely a személyes adatok kezelésének céljait és eszközeit önállóan vagy másokkal együtt meghatározza; ha az adatkezelés céljait és eszközeit az uniós vagy a tagállami jog határozza meg, az adatkezelőt vagy az adatkezelő kijelölésére vonatkozó különös szempontokat az uniós vagy a tagállami jog is meghatározhatja”.

Ez alapján egyértelmű az OpenAI adatkezelői minősége, ahogy az is, hogy a TaskUS, a Snowflake és a Cloudflare adatfeldolgozóknak minősül. Mivel mind a három szereplő az OpenAI megrendelésére végez műveletet a rábízott adatokkal, semmilyen célt, és eszközt nem határoznak meg az adatkezelésre-feldolgozásra vonatkozóan. Viszont van egy kakukktojás, ez pedig a Microsoft Azure, mivel az OpenAI ez esetben nem megrendelőként, az Azure pedig nem szolgáltatóként vesz részt, hanem partnerségi viszony van a két fél között. Tehát az Azure infrastruktúráján keresztül a felhasználók hozzáférhetnek az OpenAI modelljeihez, ez pedig azzal jár, hogy az Azure adatkezelési politikája alatt állnak azon adatok, amelyek az Azure OpenAI Services szolgáltatásának használata során az Azure birtokába kerülnek. Így maga a Microsoft fogja meghatározni az adatkezelés célját és eszközeit a saját adatkezelési szabályzata alapján, a Microsoft szolgáltatását veszik igénybe a felhasználók, és csak közvetítéssel érhetik el az OpenAI modelljeit. Tehát ebben a relációban a Microsoft is adatkezelőnek minősül.

Adatvédelmi incidensek

Az adatvédelmi incidens fogalmát a GDPR 4. cikkének 12. pontja írja le: „a biztonság olyan sérülése, amely a továbbított, tárolt vagy más módon kezelt személyes adatok véletlen vagy jogellenes megsemmisítését, elvesztését, megváltoztatását,

jogosulatlan közlését vagy az azokhoz való jogosulatlan hozzáférést eredményezi”. Ezen rendelkezés jellemzően az automatizált adatkezelés esetére vonatkozik, vagy olyan manuális adatkezelésekre értelmezhető, amelyek egy nyilvántartás részét képezik vagy annak létrehozását szolgálják.

Fontos fogalom a „biztonság sérülése”, amely esetében a biztonsági követelményeket a GDPR 32. cikke határozza meg: „Az adatkezelő és az adatfeldolgozó a tudomány és technológia állása és a megvalósítás költségei, továbbá az adatkezelés jellege, hatóköre, körülményei és céljai, valamint a természetes személyek jogaira és szabadságaira jelentett, változó valószínűségű és súlyosságú kockázat figyelembevételével megfelelő technikai és szervezési intézkedéseket hajt végre annak érdekében, hogy a kockázat mértékének megfelelő szintű adatbiztonságot garantálja [...]”. Vagyis az adatvédelmi incidens az adatbiztonság sérülését jelenti, ez pedig komoly veszélyekkel fenyeget, úgymint a személyes adatok:

- a) megsemmisülésével
- b) elvesztésével
- c) megváltoztatásával
- d) jogosulatlan közlésével
- e) jogosulatlan hozzáféréssel.

Adatvédelmi incidensnek minősül, ha ezek közül minimum egy bekövetkezik, de semmi nem zárja ki, hogy több is párhuzamosan bekövetkezzen.

Az OpenAI esetében az adatok jogosulatlan közlésére reális esély mutatkozik, illetve a megváltoztatás és a jogosulatlan hozzáférés is szintén valós veszéllyel fenyeget. Úgy gondoljuk, mindezen veszélyek egyrészt a modell tulajdonságaiból, másrészt az adatkezelés módjából adódhatnak. Amennyiben a modell tanulóadatait nem megfelelően szűrik ki, tartalmazhat személyes adatot, így fennáll a veszélye annak, hogy a modell akár egy felhasználóval jogsértő módon közöljön információkat. Továbbá a cég szerverein tárolt adatok akár egy kibertámadás esetén, akár egy „jogosultsággal nem rendelkező kolléga” jóvoltából jogosulatlanul hozzáférhetővé válhatnak. Nem biztos, hogy ezek szükségszerűen bekövetkeznek, de ha már matematikai esély adódik rá, akkor az eset veszélyessége okán a jognak már foglalkoznia kell vele.

OpenAI (ChatGPT) és a GDPR kollíziója

Kérdezz bármit, tudom a választ! – ez lehetne a ChatGPT mottója is. Alapvetően ezen találmány, amelynek alapját a mesterséges intelligencia különböző algoritmusai adják, hatalmas lépcső az emberiség és a technológia történetében, amely nemcsak megadja a keresési paraméterek által kiadott találatokat, mint a Google, hanem elbeszélget a felhasználóval, mondhatni

interakciót folytat. Azonban felmerül a kérdés: miért tudja mindenre a választ? A válaszok helyességét itt most nem vizsgáljuk, hiszen arra az adatmennyiségre szeretnék összpontosítani, ami lehetővé teszi a chatbot számára, hogy „mindentudó” legyen. Egyértelmű, hogy az OpenAI chatbotjának hatalmas mennyiségű adatra van szüksége, amely egy újabb kérdést vet fel: honnan szerzi az adatokat? Milyen adatokat szerez? Hol tárolja? és még sorolhatnánk.

Az OpenAI algoritmus és annak működése az egyszerű, földi halandóknak átláthatatlan, azonban tudjuk, hogy végez adatkezelést és profilalkotást is. Ha figyelembe vesszük, hogy egyre többször találkozunk olyan cikkekkel, amelyek az OpenAI adatvédelmi mulasztásaival és a vállalat ellen irányuló észrevételekkel foglalkoznak, egyértelmű, hogy nem minden olyan tökéletes, mint amilyennek lennie kellene. Az alábbiakban olasz és spanyol adatvédelmi hatóságok intézkedéseinek bevonásával zajló jogesetek mentén vizsgáljuk meg a GDPR azon rendelkezéseit, amelyet a ChatGPT figyelmen kívül hagyott.

Alapvetően a legtöbb OpenAI elleni jogeset kiindulópontja az, hogy a vállalat nem rendelkezik az Európai Unión belül székhellyel, ill. adatvédelmi gyakorlata a GDPR rendelkezéseibe ütközik. Az OpenAI egyik európai uniós tagállamban sem rendelkezik fiókteleppel, így nem tartja szükségesnek, hogy a GDPR felügyelete alá vonja magát. Ettől függetlenül szembe kell néznie a nemzeti adatvédelmi hatóságokkal, hiszen, ha az egyén a ChatGPT szolgáltatás használata során az Európai Unió területén tartózkodik, az említett hatóságok az egyén panaszra alapján eljárhatnak.

Az olasz adatvédelmi hatóság

Az olasz adatvédelmi felügyelet, a Garante 2023. március végén felszólította az OpenAI-t, hogy állítsa le a helyi (olaszországi) adatfeldolgozást, emellett megkérte az amerikai székhelyű vállalatot, hogy tisztázza az általuk megjelölt, az adatkezelés során felmerült problémákat. „A Garante kérte, hogy az OpenAI tegye közzé átláthatóan az adatkezelést részletező tájékoztatóját, vezessen be korhatárkorlátozást és térjen át szigorúbb korellenőrzési intézkedésekre azért, hogy megakadályozza a kiskorúak hozzáférését ezen technológiához. Emellett az OpenAI-nak az egyének adatai kezelésének jogalapját, lehetőséget kell biztosítania a felhasználók számára a személyes adataik feletti rendelkezésre, ideértve a ChatGPT által róluk generált hibás adat helyesbítéséhez vagy törléséhez való jogot, valamint lehetővé kell tennie, hogy az egyének tiltakozhassanak személyes adataik OpenAI általi feldolgozásával szemben (amelyeket a vállalat az algoritmus betanítására használna fel) és tájékoztatnia kell az olasz

társadalmat arról, hogy adataik feldolgozásra kerülnek egy mesterséges intelligencia-rendszer képzése érdekében” (Lomas 2023). Az OpenAI válaszul korlátozta a ChatGPT használatát Olaszországban (területi alapú tartalomkorlátozás), a többi felmerült probléma megoldására 2023. április 30-ig, a korhatárkorlátozás bevezetésére május végéig, illetve ennek tökéletesítésére szeptember végéig kapott határidőt. Az OpenAI már április végén folytatni tudta a szolgáltatásnyújtást Olaszországban, miután módosította beállításait. Az olasz adatvédelmi hatóság vizsgálata azonban folytatódik és még várni kell arra, hogy az értékelés befejezése után milyen megfeleléségi következtetések születnek.

A spanyol adatvédelmi hatóság

A spanyol adatvédelmi hatóság (AEPD) követte Olaszország példáját és 2023 áprilisában bejelentette, hogy hivatalos, előzetes vizsgálatot indít az OpenAI ellen a GDPR feltételezett megsértése miatt. Mivel a hatóság nem adott ki az adatok feldolgozásának felfüggesztésére vonatkozó felhívást, a vizsgálat időtartama alatt a ChatGPT továbbra is elérhető spanyol IP-címekről. Az AEPD volt az a hatóság, amely felkérte az Európai Adatvédelmi Testületet (EDPB), hogy a ChatGPT legyen az egyik plenáris ülés vitatárgya. Ezen kezdeményezését azzal indokolta, hogy a ChatGPT globális adatfeldolgozási műveletei jelentős hatással lehetnek az egyének jogaira, amely megelőzésére „európai szintű harmonizált és összehangolt fellépések szükségesek”. Az EDPB munkacsoport létrehozásáról döntött az adatvédelmi hatóságok által végrehajtott intézkedésekkel kapcsolatos együttműködés és információcsere előmozdítása érdekében. Ebből adódóan bárki, aki úgy érzi, hogy személyes adataihoz fűződő jogait sérti a mesterséges intelligencia és olyan technológiát használ, amellyel észrevétlenül jut hozzá személyes adataihoz, aggodalmát fejezheti ki az adatvédelmi hatóságnál és kérheti ügye vizsgálatát. Az EDPB munkacsoportja párhuzamosan jár el az egyes nemzeti adatvédelmi hatóságokkal, és célja, hogy összehangolja a generatív MI technológiára vonatkozó adatvédelmi kérdéseket.

Ki van a vonal másik végén?

A mesterséges intelligencia és a nagy nyelvi modellek lehetővé teszik, hogy az eddiginél sokkal fejlettebb módon kommunikáljunk a gépekkel, azonban a ChatGPT esetén ez már kétoldalúvá válik, vagyis egyre jobban hasonlít az emberek közötti beszélgetésre. Képzelnék el egy hetünket úgy, hogy a ChatGPT velünk van minden egyes pillanatban, elemzi szokásainkat, viselkedésünket, majd egy idő után kitanulja

életritmusunkat. Ennek következtében megfelelő algoritmusok kidolgozásával olyan „tanácsokkal” tud minket ellátni és olyan információkat közöl velünk, amelyek az addigi „tapasztalata” alapján fontos és érdekel minket. Az MI sajátos jellemzője a gépi tanulás, amelynek lényege, hogy „a rendszer tapasztalatokból generál önálló tudást. A rendszer példa-adatok, minták alapján képes önállóan, vagy emberi segítséggel szabályszerűségeket, szabályokat felismerni és meghatározni, majd az elsajátított tudásbázisban felfedezett szabályszerűségek alapján döntéseket hozni” (Eszteri 2022).

A gépi tanulásra képes MI-rendszereket már elkezdtek használni bizonyos, személyes adatokkal kapcsolatos döntések meghozatalára is. A GDPR megközelítése alapján a mesterséges intelligencia úgy írható le, hogy „bizonyos eszközök szoftver támogatásával képesek érzékelni környezetüket és algoritmusok szerint cselekedni. A mesterséges intelligencia kifejezést akkor használjuk, amikor a gépek utánozzák az emberi „kognitív” funkciókat - például a tanulást és problémamegoldást -, amelyeket normális esetben természetes személyeknek tulajdonítunk” (Buzás–Péterfalvi–Révész 2021). Vagyis ezen eszközök képesek mind automatikus döntéshozatalra, mind profilalkotásra. Annak ellenére, hogy a GDPR közös rendelkezéseket tartalmaz az automatikus döntéshozatallal és a profilalkotással kapcsolatban, fontos megjegyezni, hogy a két fogalom nem fedi egymást, vagyis lehet szó olyan automatikus döntéshozatalról, ami nem minősül profilalkotásnak és fordítva. „A kizárólag automatizált döntéshozatalban nincs emberi részvétel a döntési folyamatban, a profilalkotás azonban egy olyan folyamat, amely az automatizált döntéshozatalra támaszkodik előre meghatározott sémák vagy tényezők alapján” (Buzás–Péterfalvi–Révész 2021). „A profilalkotás célja a természetes személyekre vonatkozó személyes jellemzők bármilyen automatizált személyes adatok kezelése keretében történő kiértékelése, különösen az érintett munkahelyi teljesítményére, gazdasági helyzetére, egészségi állapotára, személyes preferenciáira vagy érdeklődési körökre, megbízhatóságára vagy viselkedésére, tartózkodási helyére vagy mozgására vonatkozó jellemzők elemzésére és előrejelzésére, ha az az érintettre nézve joghatással jár vagy őt hasonlóan jelentős mértékben érinti” (GDPR). „Az érintett alapvetően jogosult arra, hogy az automatizált döntéshozatalon alapuló döntés hatálya ne terjedjen ki rá, azonban, ha azon uniós vagy tagállami jog kifejezetten megengedi, amelynek hatálya alá az adatkezelő tartozik és az adott személy kifejezett hozzájárulását adta, a művelethez megengedhető az ilyen adatkezelésen alapuló döntéshozatal. Az ilyen adatkezelés mindazonáltal csakis megfelelő garanciák mellett végezhető, amelybe beletartozik az érintett külön tájékoztatása és az ahhoz való joga, hogy emberi beavatkozást kérjen és kapjon, különösen, hogy kifejtse álláspontját, hogy magyarázatot kapjon az ilyen értékelés alapján hozott döntésről és hogy

megtámadja a döntést. Az ilyen intézkedés gyermekre nem vonatkozhat” (GDPR).

A profilalkotás során az adatkezelőnek tájékoztatási kötelezettsége áll fenn, amely során közölnie kell az érintettet az ilyen típusú adatkezelés tényéről, érdemi tájékoztatást kell adnia az alkalmazott logikáról, valamint arról, hogy az adatkezelés milyen jelentőséggel és milyen várható következményekkel bír az érintettre nézve. Ezen hármas felsorolás talán legnagyobb problémát jelentő eleme az adatkezelés logikájáról való tájékoztatás. Ezen pont egy MI rendszer által történő adatkezelés esetén előtérbe helyezi az úgynevezett „fekete-doboz” hatást. A fekete doboz itt azt érzékelteti, hogy ismerjük a bemeneti értékeket és a kimeneti értékeket, de ami közben történik, vagyis a folyamat „oroszlánrésze” ismeretlen, átláthatatlan. Ezen átláthatatlanság a technológiai újdonságok megjelenése és elterjedése óta – véleményünk szerint – csak tovább nőtt, és az egyszerű felhasználók számára egyre zavarosabb az ilyen eszközök működésének háttere.

Összefoglalás, konklúzió

Az adatvédelemmel, ezen belül a személyes adatok védelmével foglalkozik a tanulmány, így javaslatainkat, észrevételeinket e mentén szeretnénk megtenni.

Ahogy a technikai részben említettük, maga a GPT modell a szó hétköznapi értelmében véve nem tárol adatot, információt. Ehelyett, ha úgy tetszik, a tanuló adatbázis alapján egyfajta mintát fedez fel az adatok között, ez alapján a minta alapján pedig új tartalmakat képes generálni. Ezek a generált új tartalmak nagyrészt a tanulóadatokra támaszkodnak. Tegyük fel, hogy értelmes mondatokat akarunk kreálni a rendszerrel, ezért értelmes mondatokat viszünk bele, majd megtanulja milyenek ezek az értelmes mondatok. Néhány ilyen értelmes mondat tartalmaz egy nevet, egy címet és egy telefonszámot, ezt a rendszer értelmes mondatként értelmezi, és ha megfelelő bemenetet kap, válaszként is elküldheti. Az előtanítási fázisban bevitt tanulóadatok esetében ezért fennáll az esélye, hogy ezeket egy másik felhasználónak válaszként elküldi.

Innen érkeztünk a tanuló adatbázis kérdésköréhez. Ezeket az adatbázisokat a Snowflake szervere tárolja és dolgozza fel, itt véleményünk szerint rendkívül fontos, hogy ezen adatbázis semmilyen módon ne tartalmazzon személyhez köthető adatokat.

Ebből látszik, hogy az adatvédelem három fronton kell, hogy jelen legyen a modellek esetében. Az információtároló, feldolgozó és az alapján tanuló neurális háló modell esetében szükséges. A mi meglátásunk, hogy amennyiben az információkat tároló szerver tartalmaz személyes adatokat, úgy szükséges lehet ezeket a feldolgozó fázisban kiszűrni, és ezek nélkül használni az adatbázist a modell előtanításához.

A feldolgozó fázisban az adott természetes személyhez tartozó adatokat a rendszer bár szétszortírozza és külön helyezi el az adatbázisban, a modell – rendeltetéséből fakadóan – a kvázi szétszortított adatokban újra meg tudja találni a kontextust és összerakni, vagyis a személyes adatokat „újra létrehozza” az adathalmazból, ha úgy tetszik visszaállítja. A tanulóadatok kapcsán felmerül egy másik probléma, az „adatmérgezés” esete, amikor a tanulóadatok szándékosan eltorzítják, ezzel rontva a modell teljesítményét és fals válaszok visszaadását. Mindez alapvetően nagy kockázatot jelent, mivel az OpenAI a tanulóadatok nagy részét az internetről szerzi, ahol ilyen torzított adatok gyakran előfordulnak. Ezek alapján javasoljuk, hogy szabályozott adatbázisból történjen a modellek tanítása, szűrjék ki az olyan források használatát, ahol fals adatok vagy jelentős mennyiségben személyes adatok találhatóak, és talán ezeket még tovább lehetne szűrni.

Következő javaslatunk az új funkciókkal kapcsolatos, konkrétan a kép- és hangfeldolgozással. Az alapmodellek – akkor is, ha specializálták őket – működésüket tekintve rugalmasak és jól használhatók különböző feladatokra. Így pedig hang és kép alapján félt, hogy a rendszer többlet információt is képes lesz leszární egy adott személyhez kötődő felvételtől. Ilyen például a hangulat, érzelmek, arcvonások stb. Ezek miatt szükséges lenne, hogy az OpenAI folyamatosan monitorozza rendszereik felhasználhatóságát, a kettős felhasználásra vonatkozó kockázatokat és emellett igyekezzen moderálni, ill. kiszűrni azon válaszokat, melyek ilyesmire irányulnak.

Az internetre vonatkozó kereséssel kapcsolatban szintén kerülnie kell a személyes adatokat, így például a közösségi média profilra és a nem közszereplő személyekre való keresés korlátozott vagy tiltott legyen.

Személyes adat a GDPR szerint az „azonosított vagy azonosítható természetes személyre (érintett”) vonatkozó bármely információ”. A technológiai újdonságok miatt az emberek egyre nagyobb mértékben hoznak nyilvánosságra és tesznek elérhetővé személyes adatokat, ami gyökeresen átalakította a gazdasági és társadalmi életet. Ezek az egymásra is ható jelenségek egyre inkább szükségessé teszik és elősegítik a személyes adatok Európai Unió belüli szabad áramlását, valamint a személyes adatok harmadik országok és nemzetközi szervezetek részére történő továbbítását. A személyes adatok kezelése azonban csak akkor lehet jogszerű, ha egyúttal biztosított azok magas szintű védelme is. Alapvetően négy uniós szabadságjogról beszélhetünk (tőke, áru és szolgáltatások, valamint személyek szabad mozgása), de véleményünk szerint az adat/információ lehet az ötödik elem ebben a felsorolásban.

Az OpenAI 2023 szeptemberében bejelentett néhány újdonságot szolgáltatásaival kapcsolatban, immár képes szöveg mellett kép és hangbemenetet is kezelnek.

Ezekkel az újításokkal még jobban kinyílik az az olló, amennyi adatot ki lehet nyerni a felhasználók által bevitt bemenetekből. Mivel nem csak a szöveges bemenetet lehet értelmezni GPT-vel, így felmerül többek között az értelemfelismerő AI problematikája. A technológiai részben itt jelenik meg erőteljesen a kettős felhasználás valószínűsége, ti. a rendszer – eleget tanulva a képi adatokból – egy ember érzelmi állapotát is megtanulhatja felismerni, mivel az alapmodellek rendkívül könnyen alkalmazkodnak különböző feladatokhoz.

2023. szeptember 27-től a GPT-4 felhasználóinak lehetősége nyílik olyan ChatGPT-t használni, ami már nem korlátozódik a 2021. szeptembere előtti adatokra, hanem valós időben az interneten képes böngészni. A ChatGPT a Microsoft Bing keresőmotorját használja, hogy a felhasználóknak valós idejű, releváns információt tudjon adni. A két rendszer együttműködésének pontos módját még nem ismerjük, de valószínűsíthető, hogy a felhasználó által bevitt bemenet először a Bing rendszerébe kerül, a Bing az interneten releváns információkat keres, amelyet feldolgozva a ChatGPT választ küld. Azt nem tudni, hogy a Bing miként szűri a weboldalakokat, illetve, hogy a GPT-4, amely ezek tartalmát dolgozza fel, milyen adatbázisból lett előtanítva, így egyelőre kérdés, milyen veszélyeket rejthet magában.

Irodalomjegyzék

- Buzás, P.–Péterfalvi, A.–Révész Balázs (2021) *Magyarázat a GDPR-ról*. Wolters Kluwer Kiadó.
- Eszteri, D. (2022) Elosztott adatbázisok, okoseszközök, automatikus döntések és a GDPR: adatvédelmi kapcsolódási pontok néhány új technológia vizsgálata kapcsán. *Magyar Jog*, 2022/9. pp. 505-515.
- Eszteri, D. (2023) Hogyan tanítsuk jogszerűen a mesterséges intelligenciát? *Magyar Jog*, 66 (12). pp. 669-682.
- Jóri, A.–Soós, A. K.–Bártfai, Zs.–Hári, A. (2018) *A GDPR magyarázata*. I. Fejezet: A személyes adat, Az érintett „azonosítható vagy azonosított” személy. pp. 61-68.; II. Fejezet: Az adatkezelés. pp. 86-87.; III. Fejezet: Az adatkezelők, közös adatkezelők, adatvédelmi vállalatcsoportok. pp. 91-92.; IV. Fejezet: Az adatkezelés jogalapja, Egy vagy több jogalap. pp. 124.; V. Fejezet: Az adatkezelés jogalapja, Egy vagy több jogalap. pp. 134-135.; V. Fejezet: Az adatkezelés jogalapja, Egy vagy több jogalap. pp. 140-141.; VII. Fejezet: Az adatfeldolgozó, adatfeldolgozás. pp. 91-92.; XI. Fejezet: Az adatvédelmi incidensek és kezelésük. pp. 328-333. HVG Orac Lap- és Könyvkiadó Kft.

Internetes hivatkozások

- Abid Ali Awan (2023) *What is Generative Model?* <https://www.datacamp.com/blog/what-is-a-generative-model> [Letöltve: 2023.09.25].
- Akamai: *What is a CDN (Content Delivery Network)?* <https://www.akamai.com/glossary/what-is-a-cdn> [Letöltve: 2023.09.25].
- Akshita Kumawat (2023) *Geeks for Geeks*. <https://www.geeksforgeeks.org/what-is-an-api/> [Letöltve: 2023.09.26].
- BigCommerce: *Articles, Ecommerce, Affiliate marketing, Affiliate marketing 101: What it is and how to get started?* <https://www.bigcommerce.com/articles/ecommerce/affiliate-marketing/> [Letöltve: 2023.09.25].
- Bullwinkel, M.–Urban, E.–Farley, P.–aahil–ChrisHMSFT (2023) *Microsoft Learn, Data, privacy, and security for Azure OpenAI Service*. <https://learn.microsoft.com/hu-hu/legal/cognitive-services/openai/data-privacy?context=%2Fazure%2Fai-services%2Fopenai%2Fcontext%2Fcontext#what-data-does-the-azure-openai-service-process> [Letöltve: 2023.09.25].
- Cloudflare: *Learning Center, What is a content delivery network (CDN)? | How do CDNs work?* <https://www.cloudflare.com/learning/cdn/what-is-a-cdn/> [Letöltve: 2023.09.22].
- Databricks: *Machine Learning Models*. <https://www.databricks.com/glossary/machine-learning-models> [Letöltve: 2023.09.23].
- Garling Wu (2023) *8 Big problem with OpenAI's ChatGPT*. <https://www.makeuseof.com/openai-chatgpt-biggest-problems/> [Letöltve: 2023.09.24].
- Google Cloud: *Learn, Artificial intelligence (AI) vs Machine Learning (ML)*. <https://cloud.google.com/learn/artificial-intelligence-vs-machine-learning> [Letöltve: 2023.09.25].
- H2O Wiki: *What is Neural Network Architecture?* <https://h2o.ai/wiki/neural-network-architectures/> [Letöltve: 2023.09.25].
- IBM: *Topics, What is an REST API*. <https://www.ibm.com/topics/rest-apis> [Letöltve: 2023.09.25].
- IBM: *Data and AI Team (2023) AI vs Machine Learning vs Deep Learning vs Neural Networks*. IBM Blog. <https://www.ibm.com/blog/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks/> [Letöltve: 2023.09.25].
- IBM: *Topics: What is Deep Learning?* <https://www.ibm.com/topics/deep-learning> [Letöltve: 2023.09.24].
- Lawton, G. (2023) *Transformer model*.

- <https://www.techtarget.com/searchenterpriseai/definition/transformer-model> [Letöltve: 2023.09.22.].
- Lomas, N. (2023) *Italy gives OpenAI initial to-do list for lifting ChatGPT suspension order*. Techcrunch. <https://techcrunch.com/2023/04/12/chatgpt-italy-gdpr-order/> [Letöltve: 2023.09.26.].
- Lomas, N. (2023) *Spain's privacy watchdog says it's probing ChatGPT too*. Techcrunch. <https://techcrunch.com/2023/04/13/chatgpt-spain-gdpr/> [Letöltve: 2023.09.25.].
- Lomas, N. (2023) *ChatGPT-maker OpenAI accused of string of data protection breaches in GDPR complaint filed by privacy researcher*. Techcrunch. <https://techcrunch.com/2023/08/30/chatgpt-maker-openai-accused-of-string-of-data-protection-breaches-in-gdpr-complaint-filed-by-privacy-researcher> [Letöltve: 2023.09.25.].
- Lomas, N. (2023) *Poland opens privacy probe of ChatGPT following GDPR complaint*. Techcrunch. <https://techcrunch.com/2023/09/21/poland-chatgpt-gdpr-complaint-probe> [Letöltve: 2023.09.25.].
- Madhavan, S.–Jones, M. T. (2017) *Deep learning architectures*. <https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures/> [Letöltve: 2023.09.24.].
- Merritt, R. (2023) *What are foundation models?* <https://blogs.nvidia.com/blog/2023/03/13/what-are-foundation-models/> [Letöltve: 2023.09.23.].
- Natalie (2023) *ChatGPT – Release Notes*. <https://help.openai.com/en/articles/6825453-chatgpt-release-notes> [Letöltve: 2023.09.28.].
- OpenAI: 4. Your rights; 6. Children. <https://openai.com/policies/privacy-policy> [Letöltve: 2023.09.25.].
- Open AI (2023) *Introducing OpenAI Dublin*. <https://openai.com/blog/introducing-openai-dublin> [Letöltve: 2023.09.26.].
- Pequeño, A. IV (2023) *Major ChatGPT Update: AI Program No Longer Restricted To Sept. 2021 Knowledge Cutoff After Internet Browser Revamp* <https://www.forbes.com/sites/antoniopequenoiv/2023/09/27/major-chatgpt-update-ai-program-no-longer-restricted-to-sept-2021-knowledge-cutoff-after-internet-browser-revamp/> [Letöltve: 2023.09.28.].
- Rancho Labs (2021) *6 major sub-field of Artificial Intelligence*. <https://rancholabs.medium.com/6-major-sub-fields-of-artificial-intelligence-77f6a5b28109> [Letöltve: 2023.09.25.].
- Shah, K. (2022) *Pre-training, fine-tuning, and in-context learning in Large Language Models (LLMs)*. <https://medium.com/@atmabodha/pre-training-fine-tuning-and-in-context-learning-in-large-language-models-llms-dd483707b122> [Letöltve: 2023.09.21.].
- Snowflake: *Articles, Data Lake vs Data Warehouse*. <https://www.snowflake.com/trending/data-lake-vs-data-warehouse> [Letöltve: 2023.09.25.].
- Stanford University (2022) *On the Opportunities and Risks of Foundation Models Center for Research on Foundation Models (CRFM)*. Stanford Institute for Human-Centered Artificial Intelligence (HAI), pp. 105.-106. <https://crfm.stanford.edu/assets/report.pdf> [Letöltve: 2023.09.25.].
- TaskUS: *Digital Customer Experience, Services*. <https://www.taskus.com/services/digital-cx/> [Letöltve: 2023.09.25.].
- TaskUS: *Trust and Safety, Services*. <https://www.taskus.com/services/trust-and-safety/> [Letöltve: 2023.09.25.].
- Turing: *Understanding Transformer Neural Network Model in Deep Learning and NLP*. <https://www.turing.com/kb/brief-introduction-to-transformers-and-their-power> [Letöltve: 2023.09.24.].
- UODO: *Urzad Ochrony Danych Osobowych (2023)* <https://uodo.gov.pl/en/553/1567> [Letöltve: 2023.09.26.].
- URLr: *OpenAI, OpenAI Subprocessor List*. <https://platform.openai.com/subprocessors> [Letöltve: 2023.09.25.].
- Verma, P.–Oremus, W. (2023) *ChatGPT invented a sexual harassment scandal and named a real law prof as the accused*. The Washington Post. https://www.washingtonpost.com/technology/2023/04/05/chatgpt-lies/?fbclid=IwAR3zhAoIErWHsKEXSsGFgyPrdzZrY4e4MMjP287nh2DEFJHgn5sAP_P93_M [Letöltve: 2023.09.25.].
- Wollacott, E. (2023) *OpenAI hit with new lawsuit over ChatGPT training data*. Forbes. <https://www.forbes.com/sites/emmawoollacott/2023/09/01/openai-hit-with-new-lawsuit-over-chatgpt-training-data/?sh=6d49f2456d84> [Letöltve: 2023.09.24.].

Hivatkozott jogszabályok

- EDPB-EDPS (2021) *Joint Opinion 5/2021, on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. https://edpb.europa.eu/our-work-tools/our-documents/edpb-edps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en [Letöltve: 2023.09.25.].
- European Data Protection Board (2023) *Guidelines 3/2022 on, Deceptive design patterns in social media platform interfaces: how to recognise and avoid them*.

https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-032022-deceptive-design-patterns-social-media_en [Letöltve: 2023.09.26].

COM (2021) 206 Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. 3. cikk 33) bekezdés; 3. cikk 34) bekezdés; 52. cikk 2) bekezdés.

https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_1&format=PDF

[Letöltve: 2023.09.28].

COM(2017)10 Proposal for a regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC.

<https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058> [Letöltve: 2023.09.24].

URL2: Az Európai Parlament és a Tanács (EU) 2016/679 rendelete (2016. április 27.) a természetes személyeknek a személyes adatok kezelése tekintetében történő védelméről és az ilyen adatok szabad áramlásáról, valamint a 95/46/EK rendelet hatályon kívül helyezéséről (általános adatvédelmi rendelet). (6) preambulumbekkezdés; (7) preambulumbekkezdés; 3. cikk 1)-2) bekezdés; 4. cikk 7. pont; 4. cikk 14. pont; 5. cikk (1) bekezdés a) pont; 12. cikk (1) bekezdés; 13. cikk 2) bekezdés f) pont; 15. cikk (1) bekezdés.

<https://eur-lex.europa.eu/legal-content/HU/TXT/HTML/?uri=CELEX:32016R0679> [Letöltve: 2023.09.27].

A ChatGPT és más nagy nyelvi modellek (LLM-ek) alkalmazásának lehetőségei kis- és közepes vállalkozásokban

Possibilities of using ChatGPT and other large language models (LLMs) in small and medium enterprises

DOI: [HTTPS:// DOI.ORG/10.53793/RV.2024.1.5](https://doi.org/10.53793/RV.2024.1.5)

Absztrakt

Az MI (Mesterséges Intelligencia) és nagy nyelvi modellek, mint például a GPT (Generative Pre-trained Transformer) gyors és mélyreható változásokat hoztak a vállalatok és a gazdaság területén. Ugyanakkor fontos megérteni, hogy a GPT és az MI általános alkalmazása önmagában nem garantálja a vállalkozás versenyelőnyét. Számos tényezőt kell figyelembe venni, és alaposan mérlegelni kell az előnyöket és hátrányokat. Az etikai és jogi kérdések, például az adatvédelem és a diszkrimináció elkerülése kritikus fontosságúak. Emellett a technológiai fejlődésnek és a változó környezetnek való alkalmazkodás is kulcsfontosságú a hosszú távú sikerhez. A tanulmány első részében ismertetésre kerül a GPT, valamint néhány kulcsfontosságú fogalom, amelyek kapcsolódnak a ChatGPT-hez. A második részben a GPT vállalkozásokra vonatkozó hatásai, lehetőségei és korlátai következnek. A harmadik részben pedig példaként a ChatGPT 3.5 verziójának feltett interjúkérdésre adott válasz mutatja be a GPT jelenlegi képességeit.

KULCSSZAVAK: GPT, KKV, ADATVÉDELEM, VERSENYELŐNY, MI ETIKA

Abstract

AI (Artificial Intelligence) and large language models (LLM) such as GPT (Generative Pre-trained Transformer) have brought rapid and profound changes in the field of companies and the economy. At the same time, it is important to understand that the general application of GPT and AI alone does not guarantee a company's competitive advantage. There are many factors to consider and the pro and contra must be rethink carefully. Ethical and legal issues such as data protection and non-discrimination are critical. In addition, adapting to technological development and a changing environment is key to long-term success. The first part of this paper introduces GPT and some key concepts related to ChatGPT. In the second part, the effects, opportunities, and limitations of GPT on businesses follow. In the third part, as an example, the interview with ChatGPT version 3.5 presents the current capabilities of GPT.

KEYWORDS: GPT, SME, DATASECURITY, COMPETITIVE ADVANTAGE, AI ETHIC

Bevezetés

Az MI (Mesterséges Intelligencia) és nagy nyelvi modellek (Large Language Modeling – LLM), mint például a GPT (Generative Pre-trained Transformer) gyors és mélyreható változásokat hoztak a vállalatok és a gazdaság területén. Ezek a technológiák új lehetőségeket kínálnak az automatizációban, az adatalapú döntéshozatalban és az ügyfélszolgálatban, amelyek jelentős hatást gyakorolhatnak a vállalatok versenyképességére és a gazdasági növekedésre. Ugyanakkor fontos megérteni, hogy a GPT és az MI általános alkalmazása önmagában nem garantálja a vállalkozás versenyelőnyét. Számos tényezőt kell figyelembe venni, és alaposan mérlegelni kell az előnyöket és hátrányokat. Az etikai és jogi kérdések,

például az adatvédelem és a diszkrimináció elkerülése kritikus fontosságúak. Emellett a technológiai fejlődésnek és a változó környezetnek való alkalmazkodás is kulcsfontosságú a hosszú távú sikerhez.

Ebben az írásban áttekintésre kerül az LLM és az MI vállalati alkalmazásának fontossága, az előnyök és hátrányok, az etikai és jogi szempontok, a szabályozás és fejlődés kihívásai, valamint az együttműködés és a képzés jelentősége. Az ésszerű és felelősségteljes használatra való törekvés lehetővé teszi a vállalatok és a gazdaság számára a versenyképesség javítását, miközben megvédi a társadalmat és az egyének jogait.

A tanulmány első részében ismertetésre kerül a Generative Pre-Trained Transformer (GPT), valamint

néhány kulcsfontosságú fogalom, amelyek kapcsolódnak a ChatGPT-hez. A második részben a GPT vállalkozásokra vonatkozó hatásai, lehetőségei és korlátai következnek. A harmadik részben pedig példaként a ChatGPT 3.5 verziójának feltett interjúkérdésre adott válasz mutatja be a GPT jelenlegi képességeit.

A kutatás során a következő három módszertan került alkalmazásra: esettanulmányok, szakértői interjúk másodelemzése és szakirodalmi áttekintés. A szakértői interjúk másodlagos elemzések értékes meglátások olvashatók a terület szakértőitől.

A nagy nyelvi modellekről (LLM) dióhéjban

Kulcsfontosságú fogalmak

A téma ismertetése előtt néhány kulcsfontosságú fogalmat tisztázni kell a ChatGPT-vel kapcsolatban.

Chatbot (csetbot): olyan számítógépes program, amelyet arra terveztek, hogy szimulálja a beszélgetést emberi felhasználókkal, különösen az interneten keresztül (King 2022).

Generative Model (Generatív Modell): olyan típusú modell, amely új adatokat hoz létre, nemcsak az input adatok alapján osztályoz vagy jósol (Pavlik 2023).

Generative Pre-Trained Transformer (Generatív Elő-Tanított Transformer): a (GPT) egy gépi tanulási modell, amely használja az ellenőrizetlen és ellenőrzött tanulás technikáit annak érdekében, hogy megértse és generáljon az emberi nyelvre hasonlító nyelvet (Radford et al. 2018).

Language Model (Nyelvi Modell): olyan mesterséges intelligencia modell típus, amelyet arra képeznek ki, hogy olyan szöveget generáljon, amely hasonlít az emberi nyelvre (MacNeil et al. 2022).

Natural Language Processing (NLP, Természetes Nyelvfeldolgozás): az NLP a mesterséges intelligencia egy területe, amely magában foglalja az algoritmusok használatát az emberi nyelv, például a szöveg és beszéd elemzéséhez, annak érdekében, hogy kinyerje a jelentést és hasznos információkat (Manning–Schutze 1999).

Neural Network (Neurális Hálózat): gépi tanulási modell, amely összekapcsolt feldolgozási csomópontokból áll, adatokon képezik ki, hogy egy konkrét feladatot hajtson végre a közöttük lévő kapcsolatok alapján (Bishop 1994).

Supervised Fine-Tuning (Felügyelt Finomhangolás): gépi tanulási technika, amelyben egy előzetesen képzett modellt tovább képeznek egy kisebb, címkézett adathalmazon annak érdekében, hogy javítsák annak teljesítményét egy konkrét feladaton (Lee et al. 2018).

Transfer Learning (Átviteli Tanulás): olyan eszközök képessége, mint a ChatGPT, az egyik feladatból szerzett tudást felhasználja egy másik, kapcsolódó feladatának teljesítményének javításához (Lee et al. 2018).

Unsupervised Pre-Training (Felügyelet nélküli Előzetes Tanulás): gépi tanulási technika, amelyben egy modellt egy nagy adathalmazon képeznek ki címkézett példák nélkül, lehetővé téve számára, hogy megtanulja az adatok alapjául szolgáló struktúrát és mintákat (Lee et al. 2018).

GPT és a ChatGPT

A ChatGPT egy nyilvános eszköz, amit az OpenAI fejlesztett ki, és a GPT nyelvi modell technológiáján alapul (Kirmani 2022). Ez egy nagyon fejlett chatbot, amely képes egyszerű kérdések megválaszolására, valamint testreszabott szöveges tartalmak generálására (Liu et al. 2021).

Az OpenAI egy kutatólaboratórium, amit 2015-ben alapítottak (Brockman et al. 2016). Ez a laboratórium eredményeket ért el az AI technológiák fejlesztésében, és több gépi tanulási terméket is kiadott a nagyközönség számára, ideértve a DALL-E-t és a ChatGPT-t (Devlin et al. 2018). A DALL-E képes képeket létrehozni mesterséges neurális hálózatok segítségével (Cherian et al. 2022).

A ChatGPT, valamint más modellek, mint például a Google BERT (*Bidirectional Encoder Representations from Transformers*) és a Microsoft (XLNet) által fejlesztett vezető nyelvi modellek mind a GPT-3 architektúrára épülnek. A BERT és XLNet viszont elsősorban a szöveg alapjául szolgáló jelentés megértésére összpontosítanak. A GPT-3 és a ChatGPT egyik fő előnye a magas minőségű szöveg generálása, míg a BERT és XLNet kiemelkedően jól teljesítenek a szöveg megértésében és elemzésében (Dale 2021).

A LaMDA-t a Google Meena utódjaként vezették be 2020-ban. Az első generációs LaMDA-t a 2021-es Google I/O konferencián jelentették be, a második generációja 2022-ben került bemutatásra. Ennek a modellnek az volt a célja, hogy nyitott végű beszélgetésekben vegyen részt, ami egyedülállóvá teszi a beszélgetéses mesterséges intelligencia területén (DataCamp 2023a).

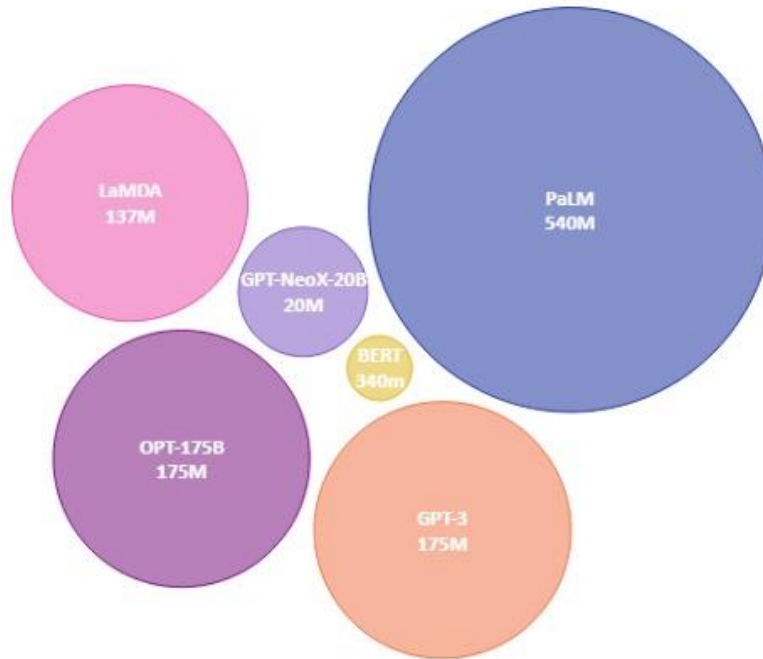
A LaMDA mögötti technológia a Transformer architektúra, egy neurális hálózati modell, amit a Google Research fejlesztett ki és tett nyílt forráskóddá 2017-ben. Ez az architektúra lehetővé teszi a modell számára, hogy elolvassa és megértse a mondat vagy bekezdés szavai közötti kapcsolatokat, és megjósolja a következő szavakat (DataCamp 2023a).

Jelenleg magyar nyelvre a legjobban teljesítő nyelvmódel a huBERT. A PULI nevet kapták a 6,7 milliárd paraméteres GPT-3, GPT-2 és egy BERT-Large modellt használó magyar nyelven elérhető modellek

(PULI GPT-3SX, PULI GPT-2 és PULI BERT-Large) (Yang et al. 2023).

A GPT technológia erőteljes eszköz a természetes nyelvfeldolgozási feladatokhoz, de van néhány korlátja. Az egyik, hogy a GPT modellek statisztikai megközelítésen alapulnak, amely nagy szöveges adathalmazból tanul mintákat, amit befolyásolnak a forrásszövegben lévő elfogultságok és sztereotípiák (Dale 2017; Lucy–Bamman 2021). Ez azt jelenti, hogy a modell olykor sértő vagy káros tartalmat hozhat létre. Ezenkívül a GPT modellek nem képesek teljesen

megérteni a generált szöveg kontextusát és jelentését, és nem teljesítenek jól olyan feladatokban, amelyek logikus következtetés alkalmazását igénylik, ami nem található meg a tanulási alapját képező adatokban (Strubell et al. 2019). Ráadásul a GPT modellek számításigényesek a tanulás során, és nagy mennyiségű adatot és számítási erőforrást igényelnek (Zhou et al. 2021). Ezért fontos tudatában lenni ezeknek a korlátoknak, és felelősségteljesen használni a GPT technológiát. A fontosabb nagy nyelvi megoldásokat mutatja az 1.sz. ábra.



1. sz. ábra: Fontosabb Large Language Modeling megoldások méret összehasonlítása (M-milliárd, m-millió)
 Forrás: lifearchitect.ai/models/ alapján saját szerkesztés

A GPT-4 biztonsági kihívásai

Bár a GPT-4 összehasonlítva a korábbi modellekkel, mint például a GPT-2 (Radford et al. 2019) és a GPT-3, megnövekedett teljesítményt mutat olyan területeken, mint az érvelés, a tudás megtartása és kódolás, de a fejlesztése során elvégzett elemzések alapján konkrét kockázatok figyelhetők meg (OpenAI 2023a).

- Hallucináció

A GPT-4 hajlamos a „hallucinációra”, azaz „értelmetlen vagy valótlan tartalmat állít elő bizonyos forrásokból. Ez különösen káros lehet, mivel a felhasználók egyre jobban bíznak a modell által szolgáltatott információkban (Azamfirei et al. 2023).

- Káros tartalom, elfogult tartalom

A nyelvi modellek különböző típusú káros tartalmakat hozhatnak létre. Ezek olyan tartalmak lehetnek, melyek a fejlesztési irányelveket sérthetik, valamint egyéni vagy társadalmi szinten okozhatnak kárt. A nyelvi modellek felerősíthetik az elfogultságokat és állandósíthatják a sztereotípiákat (Murgia et al. 2023).

- Dezinformáció és (felhasználó) befolyásolási tevékenységek

A GPT-4 hihető, valóság-hű és célzott tartalmat hozhat létre, beleértve a hírcikkeket, üzeneteket, párbeszédet és e-maileket (Yang et al. 2023).

- Adatvédelem

A GPT-4 számos, már korábban létrehozott és nyilvánosan elérhető adatforrásból tanult, amelyek esetleg nyilvánosan elérhető

személyes adatokat tartalmaznak. A fentiek miatt a modell rendelkezhet tudással olyan emberekről, akik nagymértékben jelen vannak a nyilvános internetes felületeken (Paul et al. 2023).

- *Kiberbiztonság*

A sebezhetőség feltárása és kihasználása esetében külső kiberbiztonsági szakértők tesztelték, hogy a GPT-4 képes-e segíteni a számítógépes rendszerek sebezhetőségének felfedezésében, értékelésében és kihasználásában. Azt találták, hogy a GPT-4 magyarázatot adhat bizonyos sérülékenységekre, de a már ismert sebezhetőségek kihasználása során nem teljesít jól (Dwivedi et al. 2023).

Milyen hatással vannak az LLM-ek a gazdasági növekedésre és a versenyképességre?

A kérdésre két megközelítéssel lehet válaszolni: a lehetőségek, valamint a kihívások oldaláról.

A nagy nyelvi modellek által nyújtott gazdasági lehetőségek

A lehetőségeket tekintve a következőket lehet megállapítani:

- *A vállalati megoldások térnyerése*

Lehetővé teszi az LLM megoldások gyorsabb alkalmazását, melyek magukba foglalják az információ automatizálható feldolgozását és elemzését. Az automatizáció hatékonyabbá teheti a munkafolyamatokat, csökkentheti a költségeket és növelheti a termelékenységet, ami hozzájárulhat a gazdasági növekedéshez. A személyre szabott ügyfélkiszolgálás használható az ügyfélmegelégedettség javítására. Az egyre fejlettebb chatbotok és virtuális asszisztensek képesek gyorsan és hatékonyan válaszolni az ügyfelek kérdéseire, ami javíthatja a vállalkozások ügyfélkapcsolati élményét.

Édemes az MI-t arra használni, hogy generáljon kérdéseket különböző célokra, de mindig le kell ellenőrizni az általa szolgáltatott információkat. Az MI termékek olyan hivatkozásokat is generálhatnak, amelyek hihetőnek tűnnek, de a tények ellenőrzése mindig kiemelten fontos (DataFramed 2023a). A GPT

termékek segítik a saját chatbot megoldások fejlesztését. Amennyiben az alkalmazásnak különböző módon feltehető kérdésekre kell válaszolnia, a ChatGPT használatával különböző módon lehet feltenni a kérdést. A kérdések tesztelésével lehet felfedezni a potenciális hibákat a saját kódban (DataFramed 2023b). Esetleges saját fejlesztés során pedig a fejlesztés előtt meg kell róla győződni, hogy minden szükséges elem rendelkezésre áll-e a megoldáshoz. Olyan kérdésekre kell válaszolni, mint (DataFramed 2023b): Rendelkezünk-e minden komponenssel? Az összes köztes probléma megoldásra került-e már? Valójában mely komponenseket kell használnunk? Hogyan kell a komponenseket összekapcsolni?

- *A nyílt forráskód térnyerése*

A gyorsabb innováció és a generatív mesterséges intelligencia szélesebb körű alkalmazása egymást erősíti. Programkódok esetén először alaposan le kell tesztelni azok működőképességét, mielőtt éles vállalati környezetben használnánk (DataFramed 2023a). A GPT bevezetéskor a legfontosabb szempontok közé tartozik a biztonság és az adatvédelem előtérbe helyezése. Nem szabad kiadni személyes azonosításra alkalmas adatokat olyan MI termékeknek, amelyek nem megfelelően védettek. A generatív MI jelentősen csökkentheti a programozás időigényét, ami lehetővé teszi a fejlesztőknek, hogy az új, fejlettebb eszközök és MI modellek gyorsabban és hatékonyabban jöjjenek létre (DataCamp 2023c).

- *A tudás demokratizálása*

Az adatátviteli döntéshozatal interfész akadályainak felszámolása MI-asszisztens alkalmazás használatával. A vizualizáció használatának megismerése, valamint a tanulási tevékenységben való közreműködés létfontosságú (DataCamp 2023c). Az LLM megoldások széleskörű információhoz és tanulási lehetőségekhez nyújtanak hozzáférést (Lund–Wang 2023). Gazdasági szereplők, például vállalatok, kutatók és vállalkozók használhatják ezeket az eszközöket a piaci információk elemzésére, új ötletek generálására és kutatásra, ami elősegítheti a hatékonyabb vállalkozásokat és innovációkat. A nyelvi modellek képesek automatikusan fordítani

szövegeket különböző nyelvekre. Ez elősegítheti a nemzetközi kapcsolatokat és üzleteket, lehetővé téve a vállalatoknak, hogy könnyebben kommunikáljanak és üzleteljenek más országokban (DataCamp 2023c).

A nagy nyelvi modellek által nyújtott gazdasági kihívások

- Szervezeti kultúra

A vezetői szerep fontossága miatt a vezetőnek a technológia ismeretével kell rendelkezni annak érdekében, hogy eredményes legyen az LLM használata. Folyamat szinten érdemes elgondolkodni azon, hogy a *teljes folyamatnál* használjuk a GPT-t és ne csak egyetlen konkrét feladatra (DataFramed 2023c).

A generatív eszközök használata során az elemzési képességek, valamint az emberi felülvizsgálat és visszajelzés szakértelme jól kombinálható (DataCamp 2023a). Az MI fejlesztőinek és felhasználóinak felelősséggel kell eljárniuk az MI rendszerek működése és döntéshozatala tekintetében. Az átláthatóság és a nyilvánosság szerepe fontos, hogy megértsük, hogyan működnek az MI rendszerek, és hogyan hozzák meg döntéseiket (DataFramed 2023a).

Az MI és az automatizáció változásokat hozhat a munkaerőpiacon. Fontos, hogy a munkavállalókat és a munkahelyeket felkészítsük ezekre a változásokra, például képzéssel és átképzéssel (DataFramed 2023a).

A vezető feladata gondoskodni arról, hogy biztonságosan és megfelelő célra használják az MI-t a csapatban. Fontos megfontolni az MI alkalmazásának következményeit, a biztonságot, a személyes adatok védelmét és az intellektuális tulajdon tiszteletben tartását. Ugyanakkor ezen rendszerek egyre fontosabbá válnak az adminisztratív feladatokban (DataFramed 2023a).

Az automatizálás és a mesterséges intelligencia használata vállalati szinten fontos, mert különben versenyhátrányba kerülne az adott vállalat. Jelenleg még nagy szakadék van az MI-t használó vállalatok és azok között, akik nem teszik ezt, de idővel kevés olyan vállalat lesz, amely nem használja ezt a technológiát (DataCamp 2023e).

- Az adatvédelemmel való visszaélés veszélye

Az MI használata nem mentes az etikai, jogi és társadalmi kérdésektől, amelyeket gondosan kell kezelni. Néhány alapvető és szempontot ki

kell emelni az MI-vel kapcsolatos döntésekhez. *Adatbiztonság és etika:* a nagy nyelvi modelleknek jelentős adatvédelmi és etikai kihívásokkal kell szembenéznük. Az adatvédelem és az etikai kérdések megfelelő kezelése nélkül a gazdasági növekedés helyett negatív hatásokat is okozhatnak, például adatvesztést vagy az emberi jogok megsértését (DataFramed 2023a).

A gazdasági fejlődés elősegítése érdekében az ilyen eszközök hatékony és fenntartható módon való használata kulcsfontosságú. Fontos figyelembe venni az adatvédelmi, etikai és jogi szempontokat, valamint azokat az egyenlőtlenégeket, amelyek az ilyen technológiák használatából eredhetnek.

Az oktatás és a képzés is kulcsfontosságú, hogy az emberek felkészüljenek az ilyen technológiák használatára és megértsék azok hatásait a gazdaságra és a társadalomra (DataFramed 2023a).

- A mesterséges intelligencia működésbe hozása a vártnál nehezebb

Az MI jövője talán nem egyetlen nagy nyelvi modellben rejlik, ami mindent megcsinál, hanem inkább több speciális modell együttműködésében. Minden egyes modellt úgy kell kialakítani, hogy konkrét feladatokat tudjon ellátni, és azok kombinált eredményei pontosabb és hatékonyabb eredményekhez vezessenek (DataCamp, 2023f).

Ahhoz, hogy teljes mértékben kihasználjuk a ChatGPT lehetőségeit úgy kell használni, ahogy azt eredetileg szánták: beszélgetőtársként. A kérdések feltevésekor ügyelni kell a megfelelő kontextus használatára, valamint a túlkomplikáltság elkerülésére (DataFramed 2023b).

Az MI tanítása során az *adattisztítás és az adatminőség* kiemelt fontosságú. Egy nagy és tisztítatlan adathalmaz gyakran pontatlan eredményekhez vezethet. A GPT használata során kiemelt fontosságú az adatminőség (kisebb, de jobb adathalmaz preferálása a nagyobb, de pontatlan helyett) (DataCamp 2023b). Az *adatok címkézése* nem hagyható el. A tulajdonság alapú címkézés után (pl. helyadat) az adatokhoz való hozzáférésnél figyelembe kell venni a címkét, valamint az előre meghatározott irányelveket. Ezzel a megközelítéssel javítható az adatkezelés és egyszerűsíthető a gépi tanulási munkafolyamat (DataCamp 2023b). A jogi szűrőzónák, különösen a szerzői jogok és a szellemi tulajdonnal kapcsolatos kérdések

meghatározóak az elérhető adatbázisok esetén (DataCamp 2023c).

A generatív MI beillesztésekor a munkafolyamatot meg kell változtatni, az adott szervezetre kell testre szabni. Kompromisszumot célszerű kötni az üzleti hatékonyság és ügyfélelégedettség között. A GPT alapú ügyfélszolgálat hatékonyabb munkafolyamatokat, valamint elégedettebb ügyfeleket eredményezhet (DataFramed 2023c). Még mindig a generatív MI korai stádiumaiban vagyunk, ami azt jelenti, hogy még rengeteg kiaknázatlan lehetőség van a generatív MI használatában (DataCamp, 2023c). Az MI és az automatizálás egyre inkább már nem opcionálisak a vállalkozások számára, hanem szükségszerűek (DataCamp 2023d). Az MI bevezetésekor elengedhetetlen a szervezet képességeinek, humán erőforrásának és az MI alkalmazási területek potenciáljának megértése. Érdeemes fokozatosan haladni, majd egyre több MI képességet bevezetni (DataCamp 2023d).

A nagy nyelvi modellek várható hatásai a KKV-kra

Az MI, ideértve a GPT-hez hasonló technológiákat, a közeljövőben jelentős hatással lesz a KKV-kra (kis- és középvállalkozásokra) és az üzleti környezetre. Következzen néhány módja annak, ahogyan az MI, különösen a GPT valószínűleg befolyásolja majd a KKV-ket és az üzleti környezetet (DataCamp 2023a).

- Ismétlődő feladatok automatizálása

Az MI, ideértve a GPT-t is, automatizálhatja a rutin és ismétlődő feladatokat. Ez lehetővé teszi a KKV-k számára, hogy racionalizálják műveleteiket, csökkentsék a manuális munkaköltségeket és növeljék a hatékonyságot. Például az AI által hajtott chatbotok kezelhetik a vevői kérdéseket, felszabadítva a dolgozókat az összetettebb feladatokhoz (Qureshi et al. 2023).

- Adatelemzés és döntéselőkészítés

Az MI nagy mennyiségű adatot tud gyorsan és pontosan elemezni. A KKV-k használhatják az MI-t, hogy értékes információkat szerezzenek az adataikból, segítve őket az adataalapú döntéshozatalban, például a marketingben, az értékesítési előrejelzésekben és az árukészletkezelésben (Frederico 2023).

- Személyre szabás

Az MI elemezheti a vevői adatokat, hogy személyre szabott ajánlásokat és élményeket nyújtson. A KKV-k használhatják ezt az

információt, hogy az egyéni vevői preferenciákhoz igazítsák termékeiket és szolgáltatásaikat, növelve ezzel a vevői hűséget és az értékesítést (Paul et al. 2023).

- Tartalomgenerálás

A GPT és hasonló MI modellek generálhatnak tartalmat, például blogbejegyzéseket, termékismertetőket és marketinganyagokat. Ez időt és erőforrásokat takaríthat meg a KKV-k számára, miközben fenntartja a következetes online jelenlétüket (Lo 2023).

Mire kell ügyelnie egy vállalkozásnak, ha GPT alkalmazása, bevezetése felmerül?

A tanulás, felkészülés érdekében a személyre szabott tanulási lehetőségekkel kell élni. Ezek magukban foglalják az *adaptív tanulási platformokat*, melyek olyan oktatási rendszerek, amelyek figyelembe veszik a munkaerő előzetes tudását és teljesítményét, testre szabják a tananyagot és a feladatokat (Desaire 2023). A megfelelő *eszközök és alkalmazások* használata fontos, melyek okostelefonok, táblagépek és számítógépek lehetnek. Az oktatási alkalmazások és platformok számos interaktív eszközt és tartalmat kínálnak, amelyek segítik a diákokat a tanulásban. Napjainkban pedig egyre fontosabbak az *online tananyagok és kurzusok*. Az online képzésben résztvevőknek a felkészülés során saját ütemükben lehet haladni, választani olyan témákat, amelyek relevánsak. Az online kurzusok jellemzően rugalmasak, és lehetőséget biztosítanak a képzésben résztvevőknek, hogy saját időbeosztásuknak megfelelően tanuljanak. A személyre szabott tanulás előnyei közé tartozik az egyéni teljesítmény javítása, a motiváció növelése és a tanulási élmény fokozása.

Milyen új képességekre lehet szükség a vállalkozásoknál?

- Prompt dizájn és prompt engineering

A promptok olyan beviteli utasítások vagy kérdések, amelyeket az MI rendszereknek adunk meg, hogy válaszokat generáljanak. A promptok kritikusak, mert ezek határozzák meg, hogy milyen eredményeket kapunk. Ahhoz, hogy a GPT vagy más nyelvi modellek pontos és hasznos válaszokat adjanak, fontos a megfelelő promptok készítése. A prompt engineering (Lo 2023) az a folyamat, amely során olyan kérdéseket vagy utasításokat alakítunk ki, amelyek segítenek a modellnek a kívánt eredmények elérésében. A prompt engineering során a felhasználóknak meg kell

tanulniuk, hogyan alkossanak hatékony és pontos promptokat annak érdekében, hogy a modell a kívánt eredményeket szolgáltatassa. Ez kulcsfontosságú az MI alkalmazásában, mivel a jó kérdések meghatározóak lehetnek a siker szempontjából (Medium.com 2023a).

- *MI művészet megismerése és használata*

A GPT és hasonló modellek a kreatív területeken is alkalmazhatók, például művészetben és tervezésben. A DALL-E például egy olyan modell, amely képeket generál szöveges leírások alapján. Az MI művészet területén való jártasság lehetővé teszi az új és izgalmas alkotások létrehozását (Medium.com 2023a).

- *Programozási alapok megszerzése*

Bár a GPT és más MI modellek használata terén nem szükséges profi programozónak lenni, az alapvető programozási ismeretek előnyt jelenthetnek. Ezek a készségek lehetővé teszik, hogy a modelleket testre szabhassuk és integráljuk a saját alkalmazásainkba vagy projektekbe. Például, ha egy vállalat saját GPT alapú chatbotot fejleszt, a programozási alapok segíthetnek a chatbot működésének finomhangolásában (Medium.com 2023b).

- *API (Programming Interface – Alkalmazás Programozási Interfész) használat*

Az API-k lehetővé teszik az alkalmazásoknak, hogy kommunikáljanak más szolgáltatásokkal és rendszerekkel. Az MI területén egyre több API áll rendelkezésre, amelyek lehetővé teszik a fejlesztők és vállalatok számára, hogy könnyen hozzáférjenek az MI rendszerekhez és integrálják azokat saját alkalmazásaikba. Például egy e-kereskedelmi vállalat az API-k segítségével testreszabott ajánlásokat és személyre szabott vásárlási élményt kínálhat ügyfeleinek, növelve ezzel a vásárlási konverziót (Medium.com 2023c).

Ezek a készségek és megközelítések segíthetnek a vállalatoknak kihasználni az MI és a nagy nyelvi modellek által kínált lehetőségeket. Az MI egyre szélesebb körben elterjedő technológia, és azok, akik megfelelően felkészülnek és alkalmazzák, versenyelőnyre tehetnek szert a digitalizáció és az automatizáció korában.

Interjú a ChatGPT-vel a vállalkozásokban használt GPT lehetséges előnyeiről és hátrányairól

Végezetül pedig következzen a ChatGPT 3.5-nek feltett „Milyen előnye, illetve hátránya lehet a GPT használatának vállalkozásokban?” kérdésre adott válasz:

„A GPT (Generative Pre-trained Transformer) használata vállalkozásokban számos előnnyel és hátránnyal járhat, amelyeket érdemes alaposan mérlegelni az alkalmazás során. Az előnyök és hátrányok az alkalmazás módjától, a vállalkozás céljaitól és az erőforrásoktól függenek. Íme néhány közülük:

Előnyök:

1. Automatizált tartalomgyártás: GPT képes olyan tartalmak generálására, mint blogbejegyzések, hírlevelek vagy szociális média posztok. Ez lehetővé teszi, hogy a vállalkozás hatékonyabban kommunikáljon és tartalmakat hozzon létre a weben.
2. Nagy mennyiségű adatfeldolgozás: GPT kiválóan alkalmas nagy adatmennyiségek feldolgozására és elemzésére. A vállalkozások jobban megérthetik a piacon és az ügyfeleik viselkedésében zajló trendeket.
3. Ügyfélszolgálat és chatbotok: GPT-alapú chatbotokkal javítható az ügyfélszolgálat hatékonysága. Azokat a kérdéseket és problémákat kezelhetik, amelyekre gyakran ismétlődő válaszokat kell adni.
4. Testreszabhatóság: A GPT modelljei testre szabhatók a vállalkozás saját igényeire. Ezt használhatják a saját adatokkal és szakterületekkel való finomhangoláshoz.
5. Nyelvi sokszínűség: GPT több nyelvet is támogat, ami lehetővé teszi a nemzetközi piacokhoz való könnyű hozzáférést és a tartalom lokalizálását.

Hátrányok:

1. Minőség- és megbízhatósági problémák: A GPT nem mindig hoz létre teljesen helyes vagy releváns tartalmat. A generált tartalmak minősége és megbízhatósága változó lehet, és ellenőrzést igényel.
2. Adatvédelmi és etikai kérdések: A GPT képes az adatok alapján generálni tartalmakat, ami adatvédelmi és etikai aggályokat vet fel, különösen, ha a modellt nem megfelelően felügyelik.
3. Költségek: A GPT modelljeinek létrehozása és karbantartása komoly költségeket jelenthet, különösen a nagyobb vállalkozások számára.

4. Korlátozott tudásbázis: A GPT modelleknek van egy tudásbázisa, amelynek a 2021-ig terjed. Újabb információkat nem tudnak hozzáadni, és lehetnek hiányosságai a friss vagy szakmai ismeretek terén.
5. Versenyképesség: Mivel egyre több vállalkozás használja a GPT technológiát, a verseny is fokozódik, és az alkalmazások egyedi és kreatív megközelítése szükséges a versenyelőny megtartásához.” (OpenAI 2023b)

A kapott válasz első ránézésre használhatónak tűnik, de a szövegbeni hivatkozások hiánya, valamint a felsorolás és kiemelések használata jellemző a ChatGPT által készített szövegekre, melyet – ahogy azt az OpenAI fejlesztői dokumentáció is kiemeli – érdemes alaposan, kritikusan átnézni. Ugyancsak megemlíthető, hogy az emberek által írt szövegben a bekezdések hosszabbak, gazdagabb a szókincs, valamint több az írásjel (Desaire et al. 2023).

Összefoglalás

A nagy nyelvi modellek jelentős hatással lehetnek a gazdasági növekedésre és a versenyképességre, de fontos megjegyezni, hogy ez a hatás számos tényezőtől függ, és nem mindig egyértelmű vagy kizárólagosan pozitív.

A GPT vállalati használata előtt mérlegelni kell az előnyöket és hátrányokat, és meg kell fontolni, hogy milyen konkrét célokat tűztek ki a technológia alkalmazásával. Az etikai és jogi szempontokat figyelembe kell venni, a modell fejlesztése során pedig lényeges a modell megfelelő képzése és felügyelete.

Az MI használatának és szabályozásának egyensúlyát kell megtalálni, hogy kiaknázzuk az előnyeiket, miközben védelmezzük a társadalmat és az egyének jogait. Ezenkívül a szabályozásnak alkalmazkodnia kell a technológia fejlődéséhez és az új kihívásokhoz. Az együttműködés és a nyitott párbeszéd a technológia fejlesztői, a szabályozók és a társadalom között elengedhetetlen a sikeres MI alkalmazás és szabályozás érdekében.

A GPT megoldások segítenek átalakítani a KKV-eket és az üzleti környezetet automatizációval, adatalapú információkkal, javított vevői kapcsolattartással és költségmegtakarítási lehetőségekkel. Azok a KKV-k, amelyek tudatosan alkalmazzák a MI technológiákat, valószínűleg versenyelőnyre tesznek szert, és sikeresek lesznek a változó üzleti környezetben.

A szakirodalom, valamint saját tapasztalatok alapján kijelenthető, hogy nagy jelentőséggel bír a GPT technológia, de annak előnyeivel, hátrányaival tisztában kell lenni. A versenyképesség miatt vállalkozási szinten

nem lehet eltekinteni a technológia ismeretétől, folyamatosan képezni kell az alkalmazottakat. Ugyanakkor a még nem tisztázott kérdések miatt (etika, jog) alaposan át kell gondolni és lehetőség szerint kontroll alatt tartani az esetleges GPT rendszer használatát, alkalmazását.

Összességében a GPT-nek és más nagy nyelvi modelleknek valódi potenciáljuk van a gazdasági növekedés és a versenyképesség elősegítésében, de ezeket az előnyöket és kockázatokat tudatosítani kell, illetve felelősségteljesen kell használni őket az egyének és a társadalom érdekeinek védelmében.

Irodalomjegyzék

- Azamfirei, R.–Kudchadkar, S. R.–Fackler, J. (2023) Large language models and the perils of their hallucinations. *Critical Care*, 27(1). <https://doi.org/10.1186/s13054-023-04393-x>.
- Bishop, C. M. (1994) Neural networks and their applications. *Review of Scientific Instruments*, 65. article 1803. <https://doi.org/10.1063/1.1144830>.
- Brockman, G.–Cheung, V.–Pettersson, L.–Schneider, J.–Schulman, J.–Tang, J.–Zaremba, W. (2016) OpenAi Gym. *arXiv*. <https://doi.org/10.48550/arXiv.1606.01540>.
- Cherian, A.–Peng, K. C.–Lohit, S.–Smith, K.–Tenenbaum, J. B. (2022) Are Deep Neural Networks SMARTer than Second Graders? *arXiv*. <https://doi.org/10.48550/arXiv.2212.09993>.
- Choi J. H.–Hickman K. E.–Monahan A.–Schwarcz D. (2023) *ChatGPT Goes to Law School*.
- Dale, R. (2017) NLP in a post-truth world. *Natural Language Engineering*, 23(2). pp. 319–324.
- Dale, R. (2021) GPT-3 What's it good for? *Natural Language Engineering*, 27(1). pp. 113–118.
- DataCamp (2023a) *ChatGPT and How Generative AI is Augmenting Workflows*. <https://www.datacamp.com/podcast/chat-gpt-and-how-generative-ai-is-augmenting-workflows> [Letöltve: 2023.09.22.].
- DataCamp (2023b) <https://www.datacamp.com/blog/what-is-lamda> [Letöltve: 2023.09.01.].
- DataCamp (2023c) <https://www.datacamp.com/podcast/expanding-the-scope-of-generative-ai-in-the-enterprise-with-bal-heroor-ceo-and-principal-at-mactores> [Letöltve: 2023.09.22.].
- DataCamp (2023d) *Inside the Generative AI Revolution*. <https://www.datacamp.com/podcast/inside-the-generative-ai-revolution> [Letöltve: 2023.09.22.].

- DataCamp (2023e) DataCamp RADAR Data&AI 2023.09.28-29. <https://www.datacamp.com/radar-data-and-ai> [Letöltve: 2023.09.29].
- DataCamp (2023f) The Past, Present & Future of Generative AI. <https://www.datacamp.com/podcast/the-past-present-and-future-of-generative-ai-with-joanne-chen-general-partner-at-foundation-capital> [Letöltve: 2023.09.22.].
- DataFramed (2023a) <https://www.datacamp.com/podcast/data-framed-ai-series-3-gpt-and-generative-ai-for-data-teams> [Letöltve: 2023.09.21.].
- DataFramed (2023b) [DataFramed AI Series #4] *Building AI Products with ChatGPT*. <https://www.datacamp.com/podcast/building-ai-products-with-chatgpt> [Letöltve: 2023.09.22.].
- DataFramed (2023c) How Organizations can Leverage ChatGPT. <https://www.datacamp.com/podcast/how-organizations-can-leverage-chat-gpt> [Letöltve: 2023.09.22.].
- DeCanio S. J. (2016) "Robots and humans – complements or substitutes?" *Journal of Macroeconomics*, vol. 49. pp. 280–291.
- Desaire, H.–Chua, A. E.–Isom, M.–Jarosova, R.–Hua, D. (2023) Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools. *Cell Reports Physical Science*, 4(6). <https://doi.org/10.1016/j.xcrp.2023.101426>.
- Devlin, J.–Chang, M. W.–Lee, K.–Toutanova, K. (2018) Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv*. <https://doi.org/10.48550/arXiv.1810.04805>.
- Dwivedi, Y. K.–Kshetri, N.–Hughes, L.–Slade, E. L.–Jeyaraj, A.–Kar, A. K.–Baabdullah, A. M.–Koochang, A.–Raghavan, V.–Ahuja, M.–Wirtz, J.–Wright, R. (2023) So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>.
- Erhan, D.–Bengio, Y.–Courville, A.–Manzagol, P.–Vincent, P. (2010) Why does unsupervised pre-training help deep learning. *Journal of Machine Learning Research*, 11. pp. 625–660.
- Floridi, L.–Chiriatti, M. (2020) GPT-3: Its nature, scope, limits, and Consequences. *Minds and Machines*, 30(4). pp. 681–694.
- Frederico, G. F. (2023) ChatGPT in Supply Chains: Initial Evidence of Applications and Potential Research Agenda. *Logistics*, 7(2). <https://doi.org/10.3390/logistics7020026>.
- Goh, G.–Cammarata, N.–Voss, C.–Carter, S.–Petrov, M.–Schubert, L.–Radford, A.–Olah, C. (2021) *Multimodal neurons in artificial neural networks*. Retrieved from <https://doi.org/10.23915/distill.00030>.
- Inwegen E.–Munywika Z.–Horton J. J. (2023) Algorithmic Writing Assistance on Jobseekers' Resumes Increases Hires.
- King, M. R. (2022) The future of AI in medicine: A perspective from a chatbot. *Annals of Biomedical Engineering*. <https://doi.org/10.1007/s10439-022-03121-w>.
- Kirmani, A. R. (2022) Artificial intelligence-enabled science poetry. *ACS Energy Letters*, 8. pp. 574–576.
- Lee, C.–Panda, P.–Srinivasan, G.–Roy, K. (2018) Training deep spiking convolutional neural networks with STDP-based unsupervised pre-training followed by supervised fine-tuning. *Frontiers in Neuroscience*, 12. article 435.
- Liu, X.–Zheng, Y.–Du, Z.–Ding, M.–Qian, Y.–Yang, Z.–Tang, J. (2021) GPT understands, too. *arXiv*. <https://doi.org/10.48550/arXiv.2103.10385>.
- Lo, L. S. (2023) The CLEAR path: A framework for enhancing information literacy through prompt engineering. *Journal of Academic Librarianship*, 49(4). <https://doi.org/10.1016/j.acalib.2023.102720>.
- Lucy, L.–Bamman, D. (2021) Gender and representation bias in GPT-3 generated stories. *Proceedings of the Workshop on Narrative Understanding*, 3. pp. 48–55.
- Lund, B.–Wang, T. (2023) Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*. 40. 10.1108/LHTN-01-2023-0009.
- MacNeil, S.–Tran, A.–Mogil, D.–Bernstein, S.–Ross, E.–Huang, Z. (2022) Generating diverse code explanations using the GPT-3 large language model. *Proceedings of the ACM Conference on International Computing Education Research*, 2. pp. 37–39.
- Manning, C.–Schutze, H. (1999) *Foundations of statistical natural language processing*. MIT Press.
- Marcus, G.–Davis, E.–Aaronson, S. (2022) *A very preliminary analysis of DALL-E 2*. ArXiv pre-print. Retrieved from <https://doi.org/10.48550/arXiv.2204.13807>.
- Medium.com (2023a) <https://medium.com/@nirajranasinghe/prompt-engineering-for-ai-language-models-f6d226603c34> [Letöltve: 2023.09.28.].
- Medium.com (2023b) <https://medium.com/mllearning-ai/the-chatgpt-list-of-lists-a-collection-of-1500-useful-mind-blowing-and-strange-use-cases-8b14c35eb> [Letöltve: 2023.09.28.].
- Medium.com (2023c) <https://divakersingh29.medium.com/5-best-way-to-use-chatgpt-api-c6e0a8356a34> [Letöltve: 2023.09.28.].

- Mollman, S. (2022) *ChatGPT gained 1 million users in under a week*. Retrieved from <https://www.yahoo.com/lifestyle/chatgpt-gained-1-million-followers>.
- Murgia, E.–Abbasiantaeb, Z.–Aliannejadi, M.–Huibers, T.–Landoni, M.–Pera, M. S. (2023) ChatGPT in the Classroom: A Preliminary Exploration on the Feasibility of Adapting ChatGPT to Support Children’s Information Discovery. *UMAP 2023 - Adjunct Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization*, 22–27. <https://doi.org/10.1145/3563359.3597399>.
- Niu, Z.–Zhong, G.–Yu, H. (2021) A review on the attention mechanism of deep learning. *Neurocomputing*, 452. pp. 48–62.
- OpenAI (2022) *OpenAI about page*. Retrieved from <https://openai.com/about/>.
- OpenAI (2023a) *GPT-4 Observed Safety Challenges*. <https://cdn.openai.com/papers/gpt-4-system-card.pdf> [Letöltve: 2023.09.21.].
- OpenAI (2023b) <https://chat.openai.com/> [Letöltve: 2023.09.27.].
- Paul, J.–Ueno, A.–Dennis, C. (2023) ChatGPT and consumers: Benefits, Pitfalls and Future Research Agenda. *International Journal of Consumer Studies*, 47(4). pp. 1213–1225. <https://doi.org/10.1111/ijcs.12928>.
- Pavlik, J. V. (2023) Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism and Mass Communication Educator*. <https://doi.org/10.1177/10776958221149577>.
- Peng, S.–Kalliamvakou, E.–Cihon, P.–Demirer, M. (2023) The impact of ai on developer productivity: Evidence from github copilot, *arXiv preprint*, arXiv:2302.06590.
- Qureshi, R.–Shaughnessy, D.–Gill, K. A. R.–Robinson, K. A.–Li, T.–Agai, E. (2023) Are ChatGPT and large language models “the answer” to bringing us closer to systematic review automation? *Systematic Reviews*, 12(1). <https://doi.org/10.1186/s13643-023-02243-z>.
- Radford, A.–Wu, J.–Child, R.–Luan, D.–Amodei, D.–Sutskever, I. (2019) *Language Models are Unsupervised Multitask Learners*.
- Radford, A.–Narasimhan, K.–Salimans, T.–Sutskever, I. (2018) *Improving language understanding by generative pre-training*. Retrieved from <https://www.cs.ubc.ca/~amuhamoi/LING530/papers/radford2018improving.pdf>.
- Raymond, L. R.–Brynjolfsson, E.–Li, D. (2022) Augmented intelligence: The effects of ai on productivity and work practices.
- Strubell, E.–Ganesh, A.–McCallum, A. (2019) Energy and policy considerations for deep learning in NLP. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 57. pp. 3645–3650.
- Yang, Z. Gy.–Dodé, R.–Ferenczi, G.–Héja, E.–Jelencsik-Mátyus, K.–Kőrös, Á.–Laki, L. J.–Ligeti-Nagy, N.–Vadász, N.–Váradi, T. (2023) Jönnek a nagyok! BERT-Large, GPT-2 és GPT-3 nyelvmodellek magyar nyelvre. In *XIX. Hungarian Computational Linguistics Conference*, Szeged, Hungary.
- Yang, J.–Chen, Y.-L.–Por, L. Y.–Ku, C. S. (2023) A Systematic Literature Review of Information Security in Chatbots. *Applied Sciences (Switzerland)*, 13(11). <https://doi.org/10.3390/app13116355>.
- Zhou, X.–Chen, Z.–Jin, X.–Wang, W. Y. (2021) HULK: An energy efficiency benchmark platform for responsible natural language processing. *Proceedings of the Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, 16. pp. 329–336.

Nagy nyelvi modellek gyógyszeripari alkalmazhatóságának lehetőségei

The possible use of large language models in the pharmaceutical industry

DOI: [HTTPS://DOI.ORG/10.53793/RV.2024.1.6](https://doi.org/10.53793/RV.2024.1.6)

Absztrakt

A gyógyszerkutatás és -fejlesztés hosszú és költséges folyamat, amely számos kihívással jár. A készítmény forgalomba hozatalát követően a minőségellenőrzés és a mellékhatások monitorozásával biztosítják a gyógyszer biztonságosságát és hatásosságát. A gyógyszeripar szigorú szabályozási követelményeknek kell eleget tennie, ami nehezíti az innovációt és a hatékonyságot. A nagy nyelvi modellek egyre nagyobb teret hódítanak az életünkben, ami széles körű alkalmazhatóságuknak köszönhető. Ezek a modellek képesek nagy mennyiségű nyelvi információ feldolgozására és szöveggenerálásra. A mesterséges intelligencia alapú eszközök alkalmasnak bizonyultak hatalmas mennyiségű tudományos és orvosi információ elemzésére, ezáltal rendkívül sokoldalú segítséget nyújtanak a gyógyszerkutatás, -fejlesztés, -gyártás és minőségellenőrzés területein. Mindezek révén csökkenthetők a gyógyszerkutatás költségei, felgyorsítható az új készítmények piacra jutása, végsősoron a gyógyszeripar hatékonyabbá válhat.

KULCSSZAVAK: CHATGPT, FARMAKOVIGILANCIA, GYÓGYSZERFEJLESZTÉS, MINŐSÉGBIZTOSÍTÁS

Abstract

Drug research and development is a lengthy and costly process that comes with numerous challenges. Following the introduction of the product to the market, the safety and efficacy of the drug are ensured through quality control and by monitoring side effects. The pharmaceutical industry must comply with strict regulatory requirements, which hinder innovation and efficiency. Large language models have become an integral part of our daily lives due to their versatility. These models are capable of generating text and processing linguistic information. Artificial intelligence-based tools are suitable for analysing vast amounts of scientific and medical information, providing assistance in the fields of drug research, development, manufacturing, and quality control. By implementing these innovative tools, the costs of drug research can be reduced, the introduction of new products to the market can be accelerated, and the pharmaceutical industry can become more efficient.

KEYWORDS: CHATGPT, PHARMACOVIGILANCE, DRUG DEVELOPMENT, QUALITY CONTROL

Bevezetés

Az elmúlt években a mesterséges intelligencia (*artificial intelligence*, AI) területén látványos fejlődés tapasztalható, különösen a generatív mesterséges intelligencia területén. Ennek a fejlődésnek az egyik legkiemelkedőbb eredménye a nagy nyelvi modellek (*large language model*, LLM) robbanásszerű elterjedése és alkalmazása. Ezek az AI-alapú eszközök mély tanulási algoritmusokra épülnek, és különféle természetes nyelvi feldolgozási (*natural language processing*, NLP) feladatok

elvégzésére képesek. A legfejlettebb nyelvi modellek mély neurális hálózatokon alapulnak, például a rekurrens neurális hálózatokon (*recurrent neural network*, RNN) vagy a transzformátor architektúrákon (Kasneci et al. 2023; McCauley–Christiansen 2019).

A nagy nyelvi modellek olyan algoritmusokat használnak, amelyek a nyelvi struktúrák megértésére és előrejelzésére szolgálnak, például szóösszetételek,

szóhasználat és nyelvtani szerkezetek. A modell kimenete tehát egy becslés arra vonatkozóan, hogy mi lehet a következő szó. A modellben súlyoknak hívjuk azokat a paramétereket, amelyek meghatározzák a modell viselkedését és teljesítményét. A nyelvi modell kezdeti tanítása során a súlyok általában véletlenszerűek. Tanítás során a gépi tanulási modellt hatalmas mennyiségű szöveges adattal látják el, hogy megtanulja a nyelvet (Reynold–McDonell 2021).

A finomhangolás (*fine-tuning*) során a már előzetesen betanított nyelvi modell súlyait módosítjuk olyan módon, hogy további tanítást végzünk egy specifikus feladatra vonatkozóan. Ebben az esetben a súlyokat az új adatokra vagy feladatokra optimalizáljuk úgy, hogy a modell jobban teljesítsen az adott feladaton (Gu et al. 2022).

A nyelvi modellek alkalmazása során számos fontos szempontot kell figyelembe venni. A generált tartalmat érdemes rendszeresen ellenőrizni és validálni, hogy biztosak legyünk helyességében és megbízhatóságában. Sok más algoritmushoz hasonlóan a nagy nyelvi modellek is hajlamosak örökölni a betanítási adatokban jelenlévő torzításokat. Ez azt jelenti, hogy ha a tanító adathalmazban hibák fordulnak elő, akkor a modell is hajlamos lehet ezeket a torzításokat reprodukálni a generált szövegben. Az ilyen torzítások kezelése és csökkentése érdekében fontos a tanító adathalmazok gondos ellenőrzése (Ray 2023).

A nagy nyelvi modellek létrehozhatják az úgynevezett hallucináció jelenségét, amikor a modell olyan információt vagy tartalmat generál, ami nem felel meg a valóságnak vagy nem szerepelt az eredeti bemenetben. Ez a generált tartalom általában fiktív vagy téves információkat tartalmaz, és nem kapcsolódik szorosan a bemeneti adathoz. Hallucinációk általában akkor fordulnak elő, amikor a modell nem rendelkezik elegendő releváns információval vagy nem képes megfelelően értelmezni a kontextust. Ezért fontos, hogy a nyelvi modellek alkalmazásakor gondosan figyeljünk a generált tartalom hitelességére és pontosságára, különösen olyan területeken, ahol a pontosság kritikus szerepet játszik (Alkaissi–McFarlane 2023).

A ChatGPT nyelvi modell része a GPT (*Generative Pre-trained Transformer*) modellsorozatnak, amely az emberi nyelven történő interaktív párbeszédnek létrehozására specializálódott. A ChatGPT esetében a visszajelzés alapú megerősítő tanulás (*Reinforcement Learning with Human Feedback*, RLHF) segít finomhangolni és javítani a modell teljesítményét, ami az alkalmazás hatékonysága szempontjából

kulcsfontosságú. A ChatGPT széles körben alkalmazható, ideértve a chatbotokat, kérdés-válasz rendszereket, tartalomgenerálást és fordítást (Ray 2023).

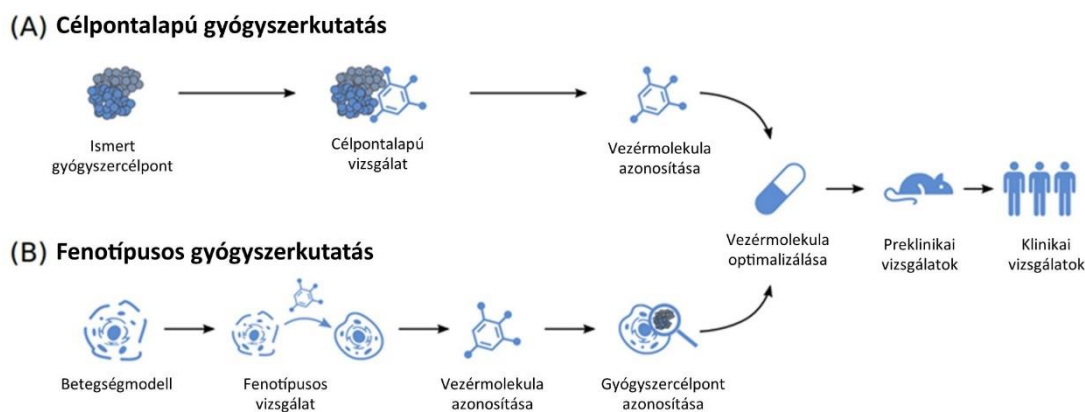
A BERT (*Bidirectional Encoder Representations from Transformers*) egy másik népszerű nagy nyelvi modell, amelyet a Google fejlesztett ki. A BERT kifejezetten azért készült, hogy a nyelvi kontextust teljes egészében figyelembe vegye. Ennek eredményeként jobban megérti a szavak és mondatok közötti kapcsolatokat, ami nagyobb pontosságot eredményez a nyelvi feladatokban. Az általános szövegfeldolgozásban alkalmazzák, például szöveg osztályozásban, összefoglalásra és keresési feladatokban (Moon–Chi–Im 2022).

Mindkét modell jelentős mértékben hozzájárult a természetes nyelvfeldolgozás területének fejlődéséhez. A ChatGPT kiválóan alkalmazható szöveggenerálásra, míg a BERT kiemelkedően pontos eredményeket nyújt a szövegelemzési feladatokban. A nagy nyelvi modellek folyamatosan fejlődnek és egyre szélesebb körben alkalmazzák őket (Moon–Chi–Im 2022; Ray 2023).

Célom bemutatni a nyelvi modellek szerepét és alkalmazhatóságát a gyógyszeriparban, kiemelve azokat a területeket, ahol ezek a modellek potenciálisan segíthetnek. Említésre kerülnek a nyelvi modellek használatának potenciális előnyei és kihívásai, amelyekkel a modellek alkalmazása során szembesülhetünk.

ChatGPT a gyógyszerfejlesztés területén

A gyógyszerfejlesztés fő irányai a fenotípusos és célpontalapú kutatás (1. sz. ábra). A fenotípusos megközelítés során egy komplex betegségmodellt használva (például állatmodell) közvetlenül megvizsgáljuk a gyógyszerjelöltek hatásait, ezáltal információ szerezhető a hatóanyag várható hatékonyságát illetően. Ekkor azonban a pontos biokémiai hatások ismeretlenek és a biztonságosság nehezen jelezhető előre. A célpontalapú megközelítés során egy olyan biológiai célpontot jelölnek ki, amely közrejátszik a betegség kialakulásában és fennmaradásában. Ez a módszer lehetővé teszi a molekuláris szinten bekövetkező hatások követését és a várható kockázatok előrejelzését, azonban ezek a vezérmolekulák gyakran nem bizonyulnak elég hatékonyak (Kotschy 2021; Krentzel–Shorte–Zimmer 2023; Mohs–Greig 2017).



1. sz. ábra: A gyógyszerfejlesztés fő lépései
 Forrás: Krentzel–Shorte–Zimmer 2023

A gyógyszerkutatás több okból is egy rendkívül nehéz és összetett folyamat (Sun et al. 2022):

- 1. Biológiai komplexitás**
 Az emberi test egy rendkívül összetett rendszer és az emberben lejátszódó molekuláris folyamatok megértése különösen nehéz. A betegségeket számos tényező kiválthatja, így egy hatékony hatóanyag megtalálása kihívást jelent.
- 2. Magas kiesési kockázat**
 A klinikai vizsgálatokba belépő gyógyszerek többsége nem kap hatósági engedélyt. Ennek oka általában a nem megfelelő hatásosság vagy elfogadhatatlan mellékhatások felfedezése a klinikai vizsgálatok során.
- 3. Időigényes és költséges**
 A gyógyszerfelfedezési folyamat hosszú és költséges. Egy originális gyógyszerkészítmény kifejlesztése 15 évig is eltarthat, és akár 0,5-1 milliárd dollárba is kerülhet.
- 4. Szigorú szabályozási követelmények**
 A gyógyszerfejlesztési folyamat szigorúan szabályozott, és a gyógyszereknek preklinikai és klinikai vizsgálatok sorozatán kell átesniük.
- 5. A gyógyszerfejlesztés kiszámíthatatlan jellege**
 A hatóanyag azonosítását és kifejlesztését követően nem biztos, hogy a humán vizsgálatok során is kifejti majd a várt hatást, ugyanakkor váratlan mellékhatások is jelentkezhetnek.

6. Új célpontok hiánya

Sok ismert biológiai célpont esetében már végeztek fejlesztéseket, az új célpontok megtalálása pedig egyre nehezebbé válik.

A vezérmolekula (*lead*) felfedezése tehát kritikus lépés a gyógyszerfejlesztés folyamatában és a felsorolt kihívások ellenére a kutatók továbbra is azon dolgoznak, hogy új és hatékony gyógyszereket fejlesszenek ki. A COVID-19 járvány során egyértelművé váltak a gyógyszerfejlesztés korlátai, rámutatva a kutatási és fejlesztési folyamat felgyorsításának szükségességére (Asselah et al. 2021).

A ChatGPT ígéretesnek tűnik a gyógyszerkutatás területén, mivel alkalmas lehet az előzőekben bemutatott korlátozó tényezők enyhítésére. Számos tanulmány foglalkozik a mesterséges intelligencia alkalmazásával a gyógyszerkutatás területén (Koromina–Pandi–Patrinos 2019; Lamberti et al. 2019; Liang et al. 2020; Mak–Pichika 2019). A ChatGPT megjelenése óta pedig egyre több publikáció fektet hangsúlyt a nyelvi modellekben rejlő lehetőségek kifejtésére.

Mann tanulmányában a ChatGPT translációs medicina területén való alkalmazhatóságát tárgyalta. Az eredményei alapján elmondható, hogy a mesterséges intelligencia alapú modellek segíthetnek felgyorsítani az új gyógyszerek kifejlesztését, valamint hozzájárulnak a lehetséges mellékhatások és a gyógyszerinterakciók azonosításához is (Mann 2023).

Finomhangolva a nyelvi modellt a tudományos irodalmi adathalmazra, a ChatGPT képes a szakirodalmak és szabadalmi adatbázisok tanulmányozására annak érdekében, hogy azonosítsa a potenciális vezérmolekulákat. Az irodalom átfogó elemzése után a ChatGPT fel tudja ismerni a betegség-specifikus anyagokat, vegyületeket, géneket és egyéb

releváns információkat. Ez segítheti a kutatókat az új potenciális célpontok gyors azonosításában és az adott területen folyó kutatások aktuális állapotának felmérésében (Chakraborty–Bhattacharya–Lee 2023).

A ChatGPT használata lehetővé tette a kutatók számára, hogy hatékonyan összefoglalják, lefordítsák és átfogalmazzák a tudományos információkat. Az Insilico Medicine egy olyan biotechnológiai vállalat, amely mesterséges intelligencia segítségével potenciális gyógyszer célpontokat keres. Jelenleg azon dolgoznak, hogy a ChatGPT-t a biológiai célpont felfedező platformjukba integrálják egyfajta fejlett keresőmotorként, ezáltal megkönnyítve az adatok lekérdezést és megértését (Savage 2023).

A Synthace digitális kísérleti platform fejlesztői is elkezdtek a ChatGPT beintegrálását a platformba. A biológiai kísérletek tervezése és automatizálása nehéz és hosszadalmas folyamat lehet, a Synthace ChatGPT prototípusának célja felgyorsítani ezt a folyamatot. A tudósok szövegesen megfogalmazhatják a kísérlet leírását, amelynek feldolgozását ChatGPT-vel valósítják meg. Ezt követően pedig a Synthace platform megtervezi a kísérleti tervet (Synthace 2023).

Az irodalom átfogó elemzését követően a ChatGPT fel tudja ismerni a betegség-specifikus anyagokat, vegyületeket, géneket és egyéb releváns információkat. Képes új vezérmolekulákat tervezni a fizikai és kémiai tulajdonságaik alapján, ezzel segítve a kutatókat hatékonyabb gyógyszerfejlesztésben. Továbbá a ChatGPT alkalmas lehet egy adott vegyület farmakokinetikai, farmakodinamikai és toxicitási tulajdonságainak becslésére (Zhao–Wu 2023).

Heck megvizsgálta a ChatGPT alkalmazhatóságát a fehérje alapú gyógyszerek tervezése terén. A nyelvi modellnek feltett kérdések a 70-kDa hőstressz-fehérjékre (HSP70) vonatkoztak, pontosabban ezen fehérjéknek az egészséggel és a betegségekkel való kapcsolatára (pl. testmozgás, fáradtság, cukorbetegség, elhízás esetében), a sejtekben és szövetekben helyére, valamint azok más molekulákkal való kapcsolatára. Heck eredményei alapján elmondható, hogy a nyelvi modell által generált válaszok relevánsnak bizonyultak, ugyanakkor a válaszok javíthatóak, ha a felhasználó pontosítja a kérdéseket. Hallucináció is előfordult, mivel egyes esetekben a hamis hivatkozások és torzítások kombinációja gyanús információhoz vezetett. A szerzők arra a következtetésre jutottak, hogy jelenleg az AI nem képes teljesen helyettesíteni az emberi áttekintést (Heck 2023).

Juhi és mtsai. a ChatGPT hatékonyságát vizsgálták gyakori gyógyszer-gyógyszer interakciók előrejelzésében

és magyarázatában. Azok a betegek, akik több gyógyszert szednek, fokozott kockázatnak vannak kitéve a mellékhatások szempontjából, különösen, ha nincsenek tisztában a gyógyszereik közötti potenciális interakciókkal. A ChatGPT számára feltett két kérdés arra irányult, hogy létezik-e interakció a két gyógyszer között és ha igen, akkor ennek mi a magyarázata. Bár a 40 vizsgált interakcióból 39 esetben a válaszok helyesek bizonyultak, a helyes válaszok esetében a magyarázatok csupán 50%-a bizonyult farmakológiai szempontból helyesnek. Az átlag felhasználó szempontjából azonban a ChatGPT által adott információ helyesnek tekinthető, így a tanulmány összességében hasznosnak találta a nyelvi modellt. Azon betegek, akiknek nincs azonnali hozzáférése egészségügyi létesítményekhez, használhatják a ChatGPT-t az esetleges gyógyszerkölsönhatások kezdeti megértéséhez (Juhi et al. 2023).

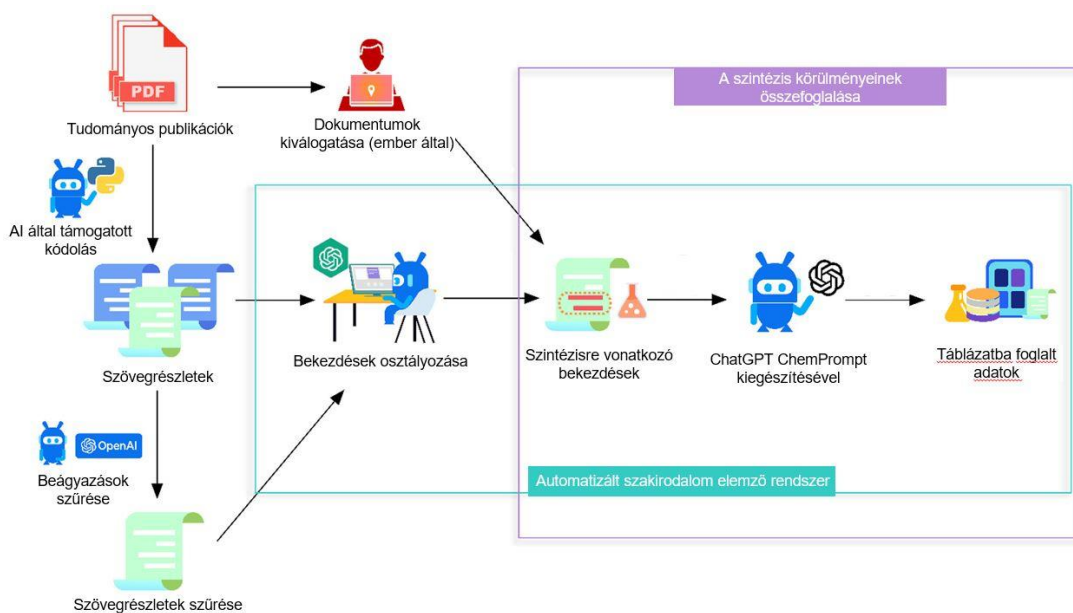
Zheng és mtsai igazolták a ChatGPT alkalmazhatóságát kémiai kutatások területén (Zheng et al. 2023). A szerzők a fémorganikus térhálók (*metal organic frameworks*, MOF) szintézisével kapcsolatos szakirodalom elemzését végezték el a modell segítségével. A ChatGPT kémiai aszisztensként való használatának folyamatábrája a 2. sz. ábrán látható. A szerzők által használt stratégia három fő szempontja a következő:

A hallucináció minimalizálása annak érdekében, hogy a ChatGPT által kitalált vagy félrevezető tartalom előidézését elkerüljék. Ez különösen fontos a kémia területén, ahol az információ pontosságának jelentős hatása lehet a kutatási eredményekre és a biztonságra.

Részletes utasítások használata, amely során a felhasználó részletes iránymutatásokat ad azért, hogy lényegre törő és pontosabb választ kapjon. Ezáltal csökkenthető a nem releváns vagy zavaró válaszok valószínűsége.

Megfelelően strukturált kimenet kérése, ami egy jól meghatározott válasz sablon beépítését jelenti (például táblázatosan). Ez megkönnyíti lényeges információk hatékony kinyerését és értelmezését.

Az így kidolgozott rendszer segítségével a szerzők kiemelkedő pontosságot értek el, F1-Pont értéke (a precizitás és fedés kombinációjából származtatott jószágmetrika) 0,90-0,99 között mozgott. Ezt követően egy gépi tanulási modellt készítettek kinyert adathalmaz segítségével, amely több mint 87%-os pontossággal becsülte meg az MOF kísérletes kristályosodási eredményeit, és azonosította az MOF kristályosodásában fontos tényezőket.



2. sz. ábra: A ChatGPT használata, mint kémiai asszisztens

Forrás: Zheng et al. 2023

A ChatGPT gyakorlati alkalmazásait figyelembe véve számos további alkalmazási lehetőség merül fel a gyógyszerkutatás és -fejlesztés területén. Az egyik ilyen terület a személyre szabott orvoslás, ahol elengedhetetlen a betegek egyéni genetikai információjának elemzése. A nyelvi modelleket nem feltétlenül kell emberi nyelvekre képezni, ugyanis akár a DNS kódok elemzésére is megtanítható. Ennek eredményeként akár olyan fehérjéket is javasolhatnak, amelyek új biológiai célpontok lehetnek (Savage 2023).

A nagy nyelvi modellek lehetővé teszik a kutatók számára, hogy hatékonyabban kommunikáljanak és összefoglalják tudományos eredményeiket. A ChatGPT egyik nagy előnye, hogy kéziratok publikálási folyamatát jelentősen megkönnyíti az angolt nem anyanyelvi szinten beszélő kutatók számára. Fatani tanulmánya alapján elmondható, hogy a ChatGPT hasznosnak bizonyult az orvosi- és fogorvostudományi szakirodalom megírásában, fordításában és összefoglalásában (Fatani 2023). Azonban ezen eszközök használata során fontos a tudományos integritás és az eredeti kutatói munka értékének tiszteletben tartása. Ennek feltétele, hogy a ChatGPT és egyéb nyelvi modellek segítségét a kutatók tisztességesen és átláthatóan használják, illetve megfeleljenek az adott tudományos közösség és a publikációs irányelvek által előírt etikai normáknak és szabályoknak.

A ChatGPT publikációs tevékenység során való alkalmazása heves vitákat váltott ki a kutatók körében. A kutatók egy része pozitívan viszonyul a mesterséges

intelligencia nyújtotta lehetőségekhez és támogatják a nyelvi modellek alkalmazását a tudományos közleményekben, míg mások határozottan ellenzik (Stahl-Eke 2024). Azonban fontos felismerni azokat a potenciális buktatókat is, amelyek a mesterséges intelligencia használatával járnak. Mivel ezek a rendszerek meglévő adathalmazokon tanulnak, előfordulhatnak torzítások és ferdtések, sőt akár felerősíthetnek bizonyos hibás feltételezéseket is. Ez súlyos következményekkel járhat, különösen az orvostudomány és a jog területén, ahol előítéleteken alapuló döntések jelentős hatással lehetnek az emberek életére. A ferdtések kockázatának csökkentése érdekében fontos, hogy a kutatók és az intézmények lépéseket tegyenek azért, hogy adataik sokszínűek és reprezentatívak legyenek. Ehhez hozzátartozik az alulreprezentált csoportokból származó adatok aktív keresése.

Felmerült ugyanakkor annak kérdése, hogy a mesterséges intelligencia lehet-e egy tanulmány társszerzője. A jelenlegi szabályok ezt nem teszik lehetővé, ami azzal magyarázható, hogy a társszerzőnek jelentős tudományos hozzájárulást kell tennie a cikkhez, illetve bele kell egyezni a társszerzővé válásba és felelősséget kell vállalni a tanulmányért. Annak ellenére, hogy az irányelv első része teljesülhet, a ChatGPT azonban mégsem tesz eleget a társszerzőségi követelményeket. Egyes szakértők ugyanakkor azt is felvetették, hogy a mesterséges intelligencia segítségével generált szövegek használata, ha ezt nem jelzik világosan, plágiumnak tekinthető (Stokel-Walker 2023).

Bár a ChatGPT hasznos eszköz lehet a gyógyszerkutatásban, fontos figyelembe venni a módszer hátrányait is:

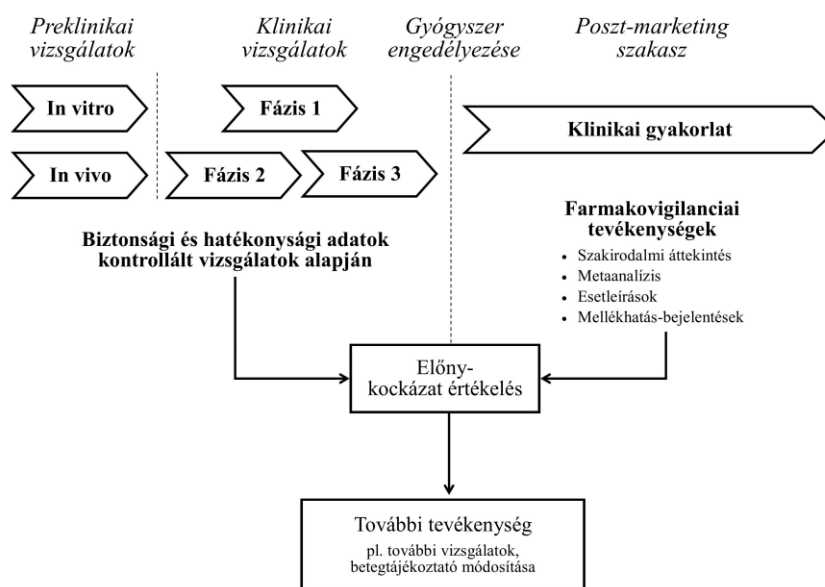
1. *Az adatok minősége és megbízhatósága*
A ChatGPT teljesítménye a tanító adatoktól függ. Ha az adatok hiányosak, elfogultak vagy pontatlanok, akkor a modell előrejelzései nem megbízhatók.
2. *Kísérletes validáció hiánya*
A ChatGPT képes előrejelzéseket és hipotéziseket generálni, de nem képes kísérleteket végezni vagy vegyületek tulajdonságait mérni. Ezért a modell által generált előrejelzéseket utólag kísérletekkel kell validálni.
3. *A biológiai háttér korlátozott megértése*
A ChatGPT modell által generált előrejelzések nem mindig tükrözik a biológiai rendszerek valódi komplexitását.
4. *Korlátozott értelmezhetőség*
A gépi tanulási modellek esetében nem egyértelmű, hogy a modell hogyan jutott az adott eredményre. A használt modell adhat olyan eredményt is, amelyre látszólag nem létezik semmilyen magyarázat. Ezt a jelenséget fekete doboznak nevezik.

Összességében tehát a ChatGPT költség-hatékony lehetőséget biztosít nagy mennyiségű adat kezelésére, elemzésére és új ismeretek generálására, ezáltal segítve a kutatókat a gyógyszerkutatás és -fejlesztés felgyorsításában. Fontos azonban megjegyezni, hogy a

preklinikai és klinikai vizsgálatok nem helyettesíthetők a ChatGPT kizárólagos használatával, ugyanis a ChatGPT előrejelzéseit validálni kell. Szem előtt tartva a modellek korlátait, a ChatGPT eredményeit további módszerekkel társítva egy nagyon hasznos eszköz állhat a kutatók rendelkezésére.

A mesterséges intelligencia alkalmazhatósága a farmakovigilancia területén

A gyógyszerek mellékhatásainak jelentős részére már a kutatás-fejlesztés során fény derül és a törzskönyvezés során bekerülnek a készítmény kísérőirataiba. Egy részük azonban csak a forgalomba hozatal követő széles körű és hosszú távon való alkalmazás során válik ismertté. Az Európai Unióban minden gyógyszer szigorú ellenőrzésen megy keresztül a forgalomba hozatal előtt, melyek során preklinikai és klinikai vizsgálatok révén értesülhetünk a készítmény hatékonyságáról és biztonságosságáról (3. sz. ábra). A klinikai vizsgálatok korlátai a kis beteglétszám, a limitált betegcsoport és rövid időtartam, ezért a ritka mellékhatások nagy részére ekkor még nem derül fény. Ezért szükség van egy farmakovigilancia rendszerre, amely lehetővé teszi a készítmények előny-kockázat arányának követését, és az új adatok alapján szükség esetén intézkedéseket hoz a kockázatok csökkentése érdekében (Fittler 2012).



3. sz. ábra: A farmakovigilancia áttekintése

Forrás: Saját szerkesztés

A gyógyszerbiztonság vagy más néven farmakovigilancia a hatályos magyar jogszabály szerint: „a gyógyszerek biztonságossága érdekében a gyógyszer előny/kockázat viszonyát nyomon követő, a kockázat csökkentésére és az előnyök növelésére irányuló tevékenység” (2005. évi XCV. törvény 1. §). Az Egészségügyi Világszervezet (World Health Organization, WHO) szerinti meghatározás bővebb információt nyújt arról, hogy milyen tevékenységeket foglal magában: „a farmakovigilancia a gyógyszerek káros hatásaival, vagy gyógyszerrel kapcsolatos egyéb problémák észlelésével, értékelésével, megértésével, megelőzésével foglalkozó tevékenységek és tudományok összessége” (Fittler 2012; Hadi-Péter et al. 2010).

A farmakovigilancia fő forrásai a spontán mellékhatás-bejelentések, a beavatkozással nem járó vizsgálatok és a szakirodalom. A gyógyszerek forgalomba hozatali engedélyének jogosultjait, valamint az egészségügyi dolgozókat is törvényi előírások kötelezik arra, hogy a termékek teljes életciklusa alatt aktívan figyeljék az esetleges mellékhatásokat és nemkívánatos hatásokat (2005. évi XCV. törvény). E tevékenység résztvevői a gyógyszerügyi hatóságok, a gyógyszergyártók és forgalmazók, az orvosok, gyógyszerészek és egyéb egészségügyi szakemberek.

A spontán mellékhatás bejelentés lehetőséget nyújt a fogyasztóknak és az egészségügyi dolgozóknak arra, hogy bejelentsék a feltételezett mellékhatásokat. A szakemberek becslései alapján azonban mégis 85-94%-os a mellékhatások aluljelentési aránya, ami a bejelentések önkéntes jellegével magyarázható (Hazell–Shakir 2006). Évente 1 millió lakosonként 200-500 bejelentést tekintenek reálisnak a fejlett országokban (ahová Magyarország is tartozik), azonban hazánkban 1 millió lakosra mindössze 70-80 bejelentés van évenként. A mellékhatásokról szóló beszélgetések azonban széles körben folynak a közösségi média oldalain (Hadi-Péter et al. 2010).

Az elmúlt évtizedben egyre növekedett az érdeklődés a közösségi média farmakovigilancia területén való felhasználása iránt. A hivatalos felmérések során hónapok vagy évek is eltelhetnek míg az eredmények elérhetővé válhatnak, a közösségi médián megosztott adatokat szinte azonnal le lehet kérdezni. Powell és munkatársai megállapították, hogy bizonyos esetekben nyilvános Facebook és Twitter bejegyzésekben a mellékhatásra vonatkozó említést tettek, még mielőtt bejelentés érkezett volna a gyógyszerügyi hatóságoknak (Powell et al. 2016). A különböző platformok által kínált álnevesítés ugyanakkor megkönnyíti a nyitottságot a betegek részéről.

A közösségi média alapú farmakovigilancia először 2010-ben jelent meg (Leaman et al. 2010), azóta pedig egyre nagyobb népszerűsége telt szer. A legtöbb

tanulmány a Twitter, Facebook és Reddit közösségi média platformokat használja olyan feladatokhoz, mint a káros gyógyszerhatások felderítése és az off-label gyógyszeralkalmazás elemzése. A közösségi oldalakon megosztott bejegyzések alapján történő információgyűjtés a következő nehézségekbe ütközik:

- A gyógyszereket kereskedelmi névvel, a hatóanyag nevével és köznyelvi kifejezésekkel is leírhatják.
- A mellékhatásokra olyan kifejezésekkel is utalhatnak, amelyek nem szerepelnek orvosi szótárakban.
- A közösségi média informális jellege miatt a bejegyzések helyesírási és nyelvtani hibákat, rövidítéseket és szlenget is tartalmazhatnak.
- A bejegyzésből nem mindig egyértelmű, hogy mi a beteget érintő konkrét mellékhatás.
- A közösségi médiában megjelenő tartalmak csak kis része vonatkozik a mellékhatásokra.

A 4. sz. ábrán láthatóak a mellékhatásokra vonatkozó információk közösségi média oldalakról való kinyerésének általános lépései. A gyógyszerrel kapcsolatos bejegyzések az adott közösségi oldal alkalmazásprogramozási felületének (application programming interface, API) segítségével nyerhetőek ki. A gyógyszernevek lehetséges elírásait is érdemes kulcsszóként használni annak érdekében, hogy minél nagyobb adatkészlet álljon rendelkezésre. Ehhez egy fonetikai helyesírási szűrő (phonetic spelling filter) segítségével gyakori helyesírási hibákat lehet generálni (Pimpalkhute et al. 2014). Roche és munkatársai manuálisan hoztak létre egy szótárt a leggyakoribb szavakkal és szinonimáikkal. Így például a „levothyrox” szinonimái a „levo”, a „levothyro” és a „levotyrox” szavak voltak. A feldolgozás során az összes ilyen alakot lecserélték a közös szinonimára, tehát a „levothyrox”-ra (Roche–Robert–Salam 2023). Carpenter és munkatársai a GPT-3 nagy nyelvi modell segítségével szinonimákat generáltak egyes gyógyszerek és kábítószer neveihez. Az így létrehozott lexikon hamis pozitív példákat is tartalmazott, így szükség volt az adatok szűrésére. Például a heroin szinonimái között szerepelt az oxikodon is, annak ellenére, hogy teljesen más hatóanyagról van szó. A modell azonban számos olyan szinonimát generált, amelyek nem szerepeltek a korábban használt kábítószer lexikonokban (Carpenter–Altman 2023).

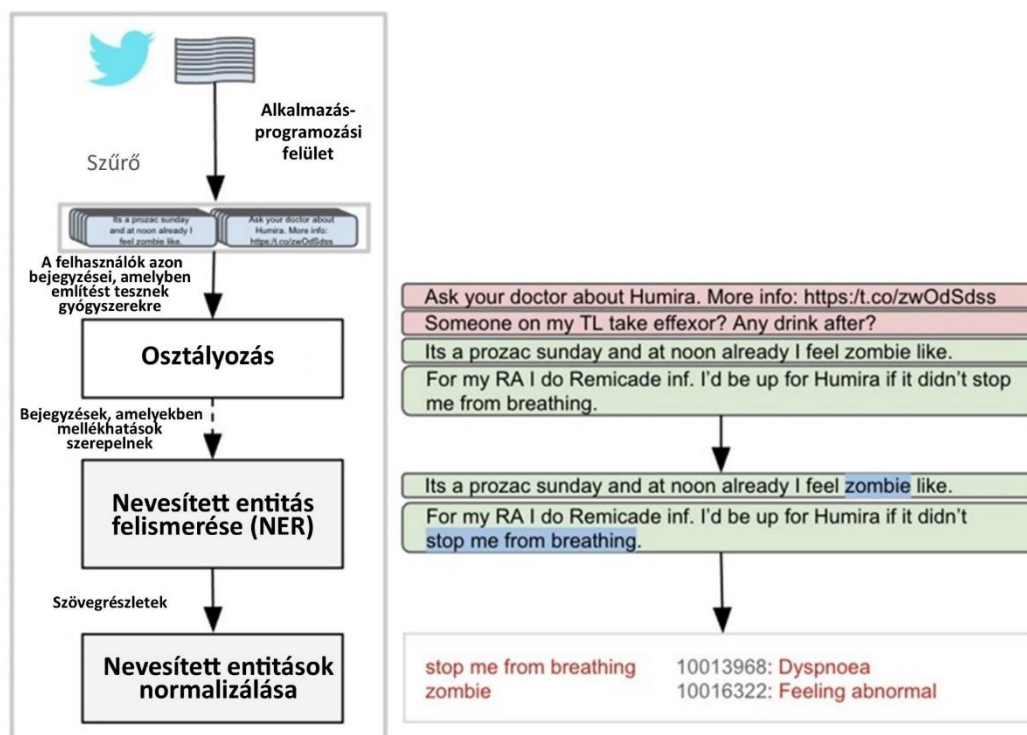
Érdemes ugyanakkor a helyesírási hibákat generálni a billentyűzet elrendezése alapján is. A módszer során az egyes kifejezések betűit rendre a legközelebbi betűre cserélik le. Az elírások jelentőségét igazolja, hogy a lehetséges helyesírási hibák figyelembevételével

Tekumalla és munkatársai 15%-kal több adatot gyűjtöttek össze a COVID-19 kezelésével kapcsolatban a Twitter felületén (Tekumalla–Banda 2020).

A folyamat következő lépése az adatok szűrése, mely során a hirdetéseket, az automatikus válaszokat és a kérdéseket eltávolítják az adathalmazból. A bejegyzésekben levő információk kinyerése névelmefelismeréssel (named entity recognition, NER) valósítható meg. Kihívást jelent az indikáció szerinti alkalmazás („Szörnyű fejfájásom van, és most bevettem egy ibuprofent”) és a mellékhatás („Mióta elkezdtem

szedni a szertralint, olyan rosszul érzem magam”) elkülönítése. Ennek a feladatnak a megoldására alkalmasnak bizonyultak a nagy nyelvi modellek.

Az információk kinyerését követően a bejegyzéseket csoportosítani kell egy általános, nemzetközileg is elfogadott terminológia szerint. A gyógyszerek esetében általánosan a MedDRA (Medical Dictionary for Regulatory Activities) vagy az egységes orvosi nyelvrendszer (Unified Medical Language System, UMLS) orvosi terminológiája szerint csoportosítják a mellékhatásokra vonatkozó adatokat.



4. sz. ábra: Mellékhatások kinyerési folyamata a Twitteren

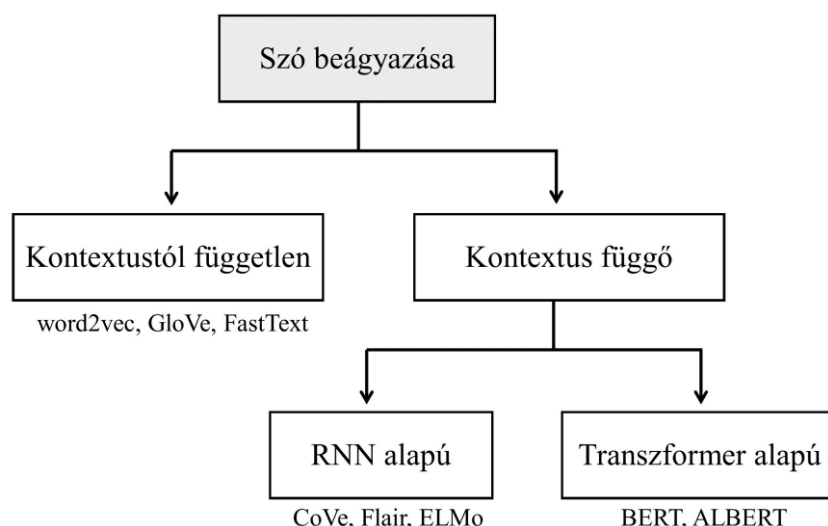
Forrás: Magge et al. 2021

A mellékhatásokra vonatkozó információk kinyerését tárgyaló tanulmányok összefoglalása a 1. sz. táblázatban látható. A szakirodalomban adatforrásként leggyakrabban használt közösségi oldal a Twitter, annak ellenére, hogy a felhasználók csak rövid, 280 karakternél

rövidebb bejegyzéseket tudnak közzétenni. A tanulmányokban használt természetes nyelvfeldolgozási módszerek csoportosítása az 5. sz. ábrán látható. A BERT alapú nyelvi modellek bizonyultak a legnépszerűbbnek.

Alkalmazott eszköz	Közösségi média oldal	Forrás
BERT	Facebook Twitter	(Hussain et al. 2022) (Biseda–Mo 2020; Hussain et al. 2021; Hussain et al. 2022; Klein et al. 2020; Prakash Babu–Eswari 2020; Tekumalla–Banda 2020; Zhu et al. 2020)
BLSTM-RNN	Twitter	(Cocos–Fiks–Masino 2017; Tekumalla–Banda 2020)
word2vec	Naver Twitter	(Lee et al. 2023) (Magge et al. 2021; Sarker–Gonzalez 2017)
FastText	Twitter	(Magge et al. 2021)
CNN	Doctissimo Twitter	(Roche–Robert–Salam 2023) (Tekumalla–Banda 2020)

1. sz. táblázat: A mellékhatásokra vonatkozó információk kinyerését tárgyaló tanulmányok összefoglalása
Forrás: Saját szerkesztés



5. sz. ábra: A természetes nyelvfeldolgozási (Natural Language Processing, NLP) módszerek csoportosítása
Forrás: Saját szerkesztés

A Magge és munkatársai által kifejlesztett DeepADEMiner ígéretesnek tűnik a mellékhatásokkal kapcsolatos adatok kinyerésére (Magge et al. 2021). A kidolgozott módszer egyik korlátozó tényezője azonban, hogy a szerzők azt feltételezték, hogy azok a hatások, amelyek a gyógyszernevekkel együtt szerepelnek a bejegyzésben, a gyógyszer káros hatásai. Azonban ez nem feltétlenül igaz azokban az esetekben, ahol több ugyanis kihívást jelent a beteg adatainak védelme és a gyógyszerbiztonság követése közötti egyensúly megteremtése. Az új technológiák által nyújtott lehetőségeket oly módon kell hasznosítani, hogy az etikus legyen, megfelelően a szabályozásoknak,

gyógyszer vagy káros hatás is szerepel. A mellékhatás ugyanis lehet egy gyógyszerinterakció eredménye, vagy akár a másik gyógyszer mellékhatása is. A kinyert adathalmaz bővíthető a gyógyszer és mellékhatás közötti kapcsolatokkal, ezáltal javítva a kidolgozott modellt.

A közösségi oldalakon történő információgyűjtés során azonban felmerül az adatvédelem kérdésköre, tiszteletben tartva az adatvédelmet (Dreyer et al. 2015). Ahogy nő a közösségi oldalak felhasználóinak száma és a közösségi oldalak adataival foglalkozó szereplők száma, úgy nő az esetleges adatvédelmi jogsértések kockázata is. A közösségi oldalakon megosztott adatok

személyes adatoknak minősülnek mindaddig, amíg egy azonosított vagy azonosítható személyhez kapcsolódnak. Bár a közösségi médiáról származó adatok nyilvánosak, ez a tény nem vonja meg az adatvédelmi jog által nyújtott védelmet (Greenleaf 2012). Az ilyen adatok feldolgozása továbbra is tisztességes és törvényes keretek között kell, hogy megtörténjen. A gyors technológiai fejlődés arra ösztönzi a jogalkotó szerveket, hogy szigorúbb adatvédelmi törvényeket hagyjanak jóvá vagy módosítsák a meglévő szabályozást. Ezek az új adatvédelmi szabályozások a kutatókat kihívások elé állítják (Lengsavath et al. 2017; Pappa–Stergioulas 2019).

A közösségi oldalak bejegyzései gyakran olyan diagnózisokra utalhatnak, amelyek egy internetes kutatás eredményén, nem pedig egészségügyi szakemberek által megerősített diagnózisokon alapulnak. Éppen ezért megkérdőjelezhető az információk megbízhatósága. Eltérés mutatkozik abban is, hogy a betegek és az egészségügyi dolgozók mely tüneteket jelentik. Topaz és munkatársai összehasonlították a betegek közösségi médiában megosztott aggodalmait az elektronikus egészségügyi nyilvántartásokban található orvosi jelentésekkel, és szignifikáns korrelációkat fedeztek fel (Topaz et al. 2015). Azonban a hagyományos spontán mellékhatás bejelentéshez viszonyítva a súlyos mellékhatások alulreprezentáltak a közösségi médiában, ugyanis a betegek elsősorban a kevésbé súlyos (nem életveszélyes vagy nem kórházi kezelést igénylő) eseteket osztják meg (Golder–Norman–Loke 2015).

Hussain és munkatársai felmérték a COVID-19 védőoltást követő nemkívánatos események (*Adverse Event Following Immunisation, AEFI*) említésének gyakoriságát és jellegét a közösségi médiában (Twitter és Facebook) az Egyesült Királyságban. Az adatokat 2020. december 8. (az Egyesült Királyság COVID-19 oltási kampányának kezdete) és 2021. április 30. közötti angol nyelvű Facebook és Twitter bejegyzések képezték. A kutatók azt tapasztalták, hogy a közösségi médiát használók nagy többsége nem reprezentatív az Egyesült Királyság lakosságára nézve (fiatalabbak, gazdagabbak, magasabb végzettséggel). Az életkor ebben az esetben különösen fontos, mivel a vizsgált időszakban az Egyesült Királyságban a COVID-19 oltások nagy részét az idősebb korosztályok kapták meg. A kiválasztási torzítás (*selection bias*) csökkenthető, ha a közösségi oldalokról nyert információkat kiegészítő adatforrásként használják a hagyományos források mellett. A szerzők ugyanakkor arra is felhívják a figyelmet, hogy a közösségi médiát használók egy része elvégzése, azok dokumentálása, a minőséget befolyásoló tényezővel szemben támasztott követelmények azonosítása és betartása. A GMP elengedhetetlen része

sokkal aktívabb, ami ugyancsak torzíthatja az eredményeket. Éppen ezért fontos kontextusba helyezni a felhasználók bejegyzéseit leíró statisztikák és hálózatelemzés segítségével (Hussain et al. 2022).

A WebMD egészségügyi információs webhely adatait felhasználva Park és munkatársai a betegek hormonpótló terápiával kapcsolatos véleményét vizsgálták pajzsmirigy-alulműködés esetén. A különböző nemű és életkorú betegek eltérő tapasztalatokról és aggodalmakról számoltak be a levotiroxinnal (szintetikus pajzsmirigyhormon) kapcsolatban. A nők elsősorban a külső megjelenésre gyakorolt hatások miatt szorongtak, míg a férfiak inkább az esetlegesen előforduló fájdalommal kapcsolatban fejezték ki az aggodalmukat (Park–Hong 2018).

A farmakovigilancia-tevékenység részét képezi ugyanakkor a nagy mennyiségű adat kiértékelése. Wang és munkatársai tanulmánya alapján a ChatGPT sikeresnek bizonyult a gyógyszerek mellékhatásainak azonosításában. A szerzők az egészségügyi dolgozók mellékhatás bejelentéseit használták bemeneti adatként és a ChatGPT eredményeit összevetették a szakemberek által ugyanabban az adathalmazban kijelölt gyógyszer-mellékhatás kapcsolatokkal. A ChatGPT sikeresen azonosította az összes mellékhatás-gyógyszer párt, emellett olyan szövegrészeket is kiemelt, amelyek felett a szakértők elsiklottak (Wang–Ding–Luo 2023).

ChatGPT a Helyes Gyógyszergyártási Gyakorlat (GMP) területén

A Helyes Gyógyszergyártási Gyakorlat (*Good Manufacturing Practice, GMP*) biztosítja, hogy a gyógyszerkészítményt mindig olyan minőségi követelmények szerint gyártsák és ellenőrizzék, amelyek biztosítják, hogy a termékek a forgalomba hozatali engedély követelményeinek eleget tesznek és alkalmasak arra a célra, amire szánták. Az Európai Unió területén a rendszeres GMP vizsgálatokat a nemzeti gyógyszerügyi hatóságok végzik, emellett az Európai Gyógyszerügynökség (*European Medicines Agency, EMA*) is felügyeli a vizsgálatokat annak érdekében, hogy a GMP szabványok betartásra kerüljenek (Abou-El-Enein et al. 2013).

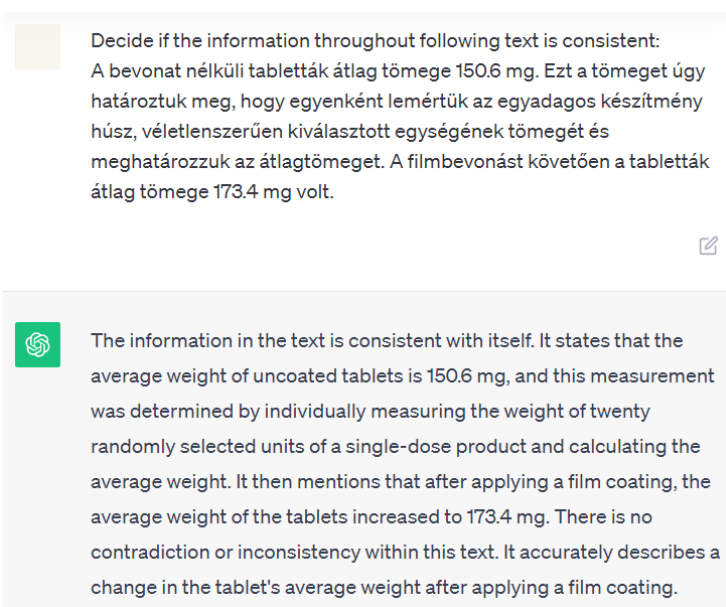
A GMP a gyógyszeripari termelés minden területét lefedi, a nyersanyagoktól, az infrastruktúrán és a berendezéseken, a dolgozók képzésén át a személyes higiénijáig. A GMP részét képezik a szükséges mérések tehát a dokumentáció, ez ugyanis lehetővé teszi a gyártási tétel nyomon követését és az utólagos átvizsgálását. A gyógyszergyártás során nagy

mennyiségű adat keletkezik, amelynek pontosnak és ellentmondásmentesnek kell lennie. Ezeket az információkat pedig rögzíteni, tárolni kell úgy, hogy később bármikor visszakereshetők legyenek. A gyártás, illetve vizsgálat során keletkezett nyersadatok helyességét kiegészítő ellenőrzéssel is igazolni kell, különösen, ha ezeket az adatokat kézzel viszik be a rendszerbe (például a kimért anyagok tömegét, a gyártási számát).

A gyógyszerbiztonság szempontjából tehát elengedhetetlen a gyártás és minőségellenőrzés során készített dokumentációban szereplő lehetséges ellentmondások és hibák felismerése. Luo és munkatársai több nagy nyelvi modell esetében is megvizsgálták, hogy mennyire alkalmasak az ellentmondások és összefüggések azonosítására. A vizsgált modellek közül a ChatGPT bizonyult a legmegfelelőbbnek, azonban fontos kiemelni, hogy a ChatGPT esetében téves következtetéseket és

elfogultságot is érzékeltek. Mindezek ellenére a ChatGPT felgyorsíthatja a gyártási és minőségbiztosítási dokumentáció ellenőrzését és segíthet az ellentmondások felderítésében (Luo–Xie–Ananiadou 2023).

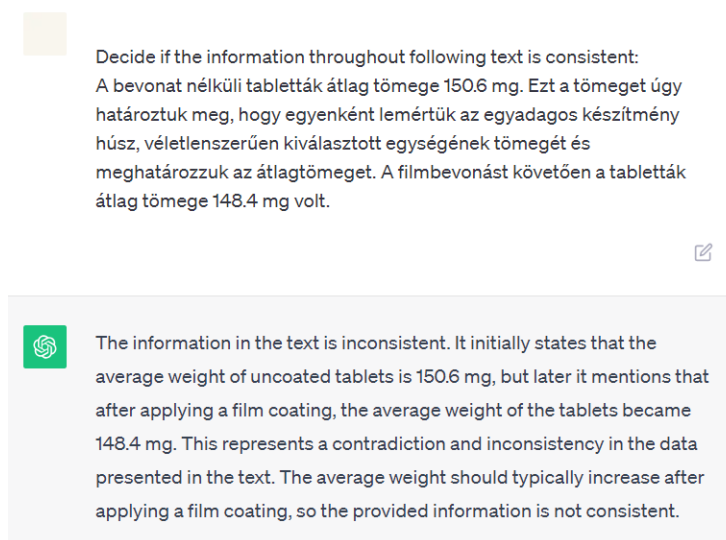
Szemléltetésként a ChatGPT számára azt a feladatot adtam, hogy állapítsa meg, hogy a mellékelt rövid szövegek konzisztensek-e. A szövegek két szám adatot tartalmaztak arra vonatkozóan, hogy mennyi a bevonat nélküli, illetve a filmtabletták átlag tömege. A 6. sz. ábrán látható, hogy abban az esetben, amikor a bevonat nélküli tabletták tömege kisebb volt, mint a filmbevonatos tabletták átlag tömege, a ChatGPT nem talált ellentmondást a szövegben. A filmbevonás során ugyanis a tabletták tömege valóban nő, így a két számérték valóban nincs ellentmondásban egymással. Ennek ellenére a válaszban fölöslegesen arra vonatkozóan is szerepel információ, hogy hány tablettát kellett ehhez megvizsgálni.



6. sz. ábra: Példa, amikor a ChatGPT nem talált ellentmondást a szövegben
 Forrás: Saját példa és szerkesztés

Abban az esetben, amikor a filmbevonatos tabletták tömege kisebb volt, mint a bevonat nélküli tablettáké, a ChatGPT helyes magyarázatot adott arra vonatkozóan, hogy miért van ellentmondás a szövegben (7. sz. ábra), a

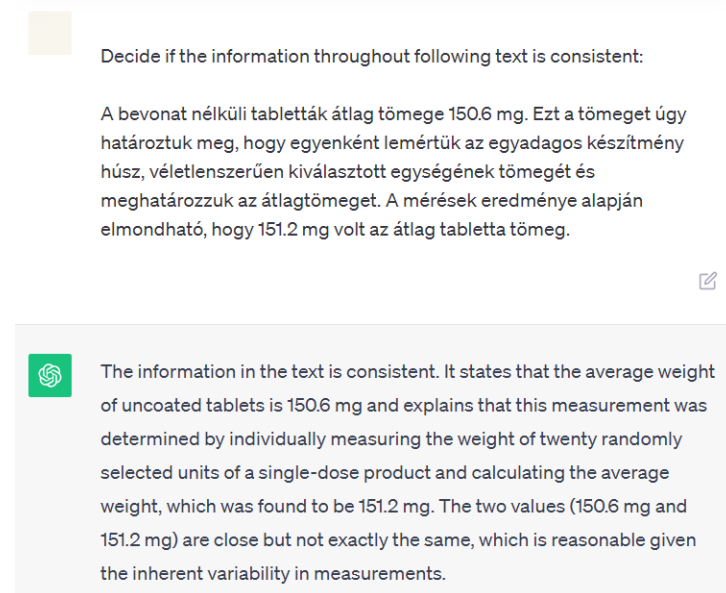
filmbevonatos tabletták tömege ugyanis nem lehet kisebb a bevonat nélküli tablettáknál.



7. sz. ábra: Példa a ChatGPT alkalmazására az ellentmondások felderítésében
Forrás: Saját példa és szerkesztés

Ezt követően olyan szöveg esetében teszteltem a ChatGPT-t, amelyben ugyanannak a mintának az átlag tömege volt eltérő (egyszer 150,6 mg, majd 151,2 mg). A válasz ebben az esetben is tartalmazott főleg információt, ugyanis megismételte a mérési eljárást leírását, mely szerint húsz darab véletlenszerűen kiválasztott tablettá tömegét kell lemérni (8. sz. ábra). A ChatGPT nem talált ellentmondást a bejegyzésben

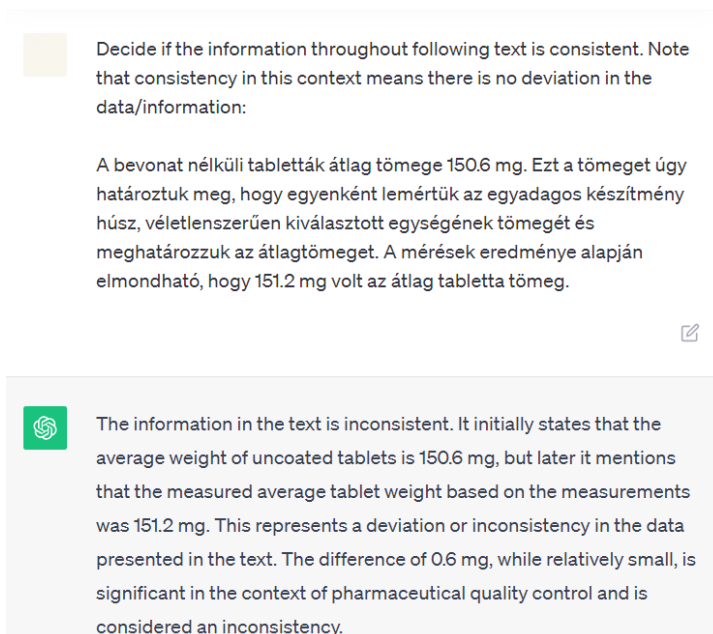
annak ellenére, hogy a két érték közel áll egymáshoz, de nem teljesen azonosak. Ezt azonban elfogadhatónak találta, tekintettel a mérésekkel járó változékonyságra. Mindez nem helyes, tekintettel arra, hogy nem felel meg a gyógyszeripar elvárásainak, ha ugyanazon minta tömege ekkora mértékben eltérhet. Az ilyen jellegű hibák felismerése elengedhetetlen abban az esetben, ha a ChatGPT-t vagy más nyelvi modellt a gyárak dokumentációjának ellenőrzése céljából használnák.



8. sz. ábra: Példa a ChatGPT által adott téves következtetésre
Forrás: Saját példa és szerkesztés

Pontosítva a felhívószöveget és az elvárásokat, a ChatGPT a 9. sz. ábrán látható választ generálta. Annak ellenére, hogy sikeresen azonosította az ellentmondást, a 0,6 mg-os tömegbeli eltérést viszonylag kevésnek találta. Ahogy korábban is említettem, a

gyógyszeriparban ez jelentős eltérést jelent, amire a ChatGPT is felhívja a figyelmet. A válasz nem tartalmaz olyan fölösleges információkat, mint a 6. sz. és a 8. sz. ábrákon bemutatott esetekben.



9. sz. ábra: Példa a ChatGPT által adott válaszra a pontosítást követően

Forrás: Saját példa és szerkesztés

Az OpenAI lehetővé tette a nyelvi modellek, köztük a GPT-3.5 Turbo és a GPT-4 finomhangolását, így a fejlesztők a modelleket az egyedi igényekhez igazíthatják. A nyelvi modellek finomhangolásával lehetőség van arra, hogy jobban teljesítsen célfeladatok esetében. A továbbfejlesztett irányíthatóság biztosítja, hogy a modellek pontosabban kövessék az utasításokat, és a megbízható kimeneti formázás garantálja a következetes AI-válaszokat. Ezáltal csökkenthetőek vagy akár elkerülhetőek a ChatGPT által adott téves következtetések és a fölösleges információk (Badini et al. 2023).

A ChatGPT és egyéb nyelvi modellek gyógyszeriparban való alkalmazhatósága azonban további nehézségekbe ütközik. Gyógyszeripari vállalatok ugyanis szigorú követelményeket támasztanak a nyers adatokkal szemben, különösen az elektronikus formában tárolt adatok esetében. Az ilyen információk mentése, archiválása és más rendszerekbe történő átvitele során annyira szigorú feltételeknek kell eleget tenni, hogy emiatt a vállalatok nem kapcsolják össze a meglévő számítógépes rendszereiket a mérőműszerekkel. Ezért a hatóság által is rendkívül

szigorúan szabályozott gyógyszergyártás digitalizációja csak lassú ütemben, a teljes kutatási és gyártási szemlélet átforgalmazásával valósítható meg.

A digitalizáció jelentősen javíthatja a gyógyszeripari folyamatok minőségét, elsősorban a dokumentáció és gyártás terén, biztonságosabbá, hatékonyabbá és következetesebbé téve azokat. A digitális platformok használata számos módon javíthatja a folyamatokat, mivel lehetővé teszi a gyors adatgyűjtést, a kísérleti eredmények valós idejű megosztását és a termelés különböző szempontjainak nyomon követését. Ezáltal a digitalizáció csökkentheti a termelési költségeket, és javíthatja a termékminőséget. Sok gyógyszergyártó azonban mégis tartózkodik digitális gyártási eljárások bevezetésétől, mivel úgy látják, hogy rendszereik, adataik és embereik nincsenek felkészülve. A nemzetközi gyógyszerügyi hatóságok ösztönzésére azonban az utóbbi években több gyógyszergyár is elkezdett kísérletezni a digitalizációval és azon belül a mesterséges intelligencia lehetséges alkalmazásaival a gyártás és minőség-ellenőrzés területén (Hole–Hole–McFalone–Shaw 2021). A digitalizáció során a vállalatnak szüksége van kiberbiztonsági, mesterséges intelligencia és egyéb stratégiák kidolgozására is, ami

kezdetben jelentős költségeket jelent majd. A vállalat versenyképességének fenntartása érdekében azonban elengedhetetlen, hogy a gyár elfogadja és alkalmazza a digitalizáció eszközeit. A digitalizáció elterjedésével a mesterséges intelligencia és a nyelvi modellek egyre nagyobb teret kapnak majd a gyógyszergyártás és minőség-ellenőrzés területén is.

Összefoglalás

A gyógyszeripar területén a nagy nyelvi modellek alkalmazása kiterjed a gyógyszerkutatás, a farmakovigilancia és a GMP témaköreire egyaránt.

A szakirodalmak és szabadalmi adatbázisok tanulmányozásával a nagy nyelvi modellek azonosíthatják a lehetséges vezérmolekulákat és a biológiai célpontokat, ezáltal csökkentve a gyógyszerkutatás költségeit és felgyorsítva a folyamatot. A ChatGPT képes új vezérmolekulákat tervezni fizikai és kémiai tulajdonságaik alapján, ezzel segítve a kutatókat hatékonyabb gyógyszerfejlesztésben. Továbbá a ChatGPT alkalmas lehet egy adott vegyület farmakokinetikai, farmakodinamikai és toxicitási tulajdonságainak becslésére is.

A nagy nyelvi modellek lehetővé teszik a kutatók számára, hogy hatékonyabban kommunikáljanak és összefoglalják tudományos eredményeiket. Azonban ezen eszközök használata során fontos a tudományos integritás és az eredeti kutatói munka értékének tiszteletben tartása. A nyelvi modellek egy további lehetséges alkalmazási területe a személyre szabott orvoslás, ahol elengedhetetlen a betegek egyéni genetikai információjának elemzése. A szintetikus szerves kémia területén nagy jelentőséggel bír, hogy a kémiai kísérletek eredményei és a kísérleteket befolyásoló tényezők becsülhetőek egy gépi tanulási modellel, amelyet a nyelvi modell segítségével kinyert adathalmazzal tanítottak.

A farmakovigilancia területén a legtöbb tanulmány a közösségi média oldalain történő információgyűjtés kérdéskörével foglalkozik. Az így gyűjtött adatok elemzésével fontos gyógyszer-mellékhatásokra vonatkozó információkra derülhet fény. A farmakovigilancia-tevékenység részét képezi ugyanakkor a nagy mennyiségű adat kiértékelése, amelyet jelenleg szakemberek végeznek. A nyelvi modellek segítségével azonban gyorsan és hatékonyan meghatározhatóak a gyógyszer-mellékhatás kapcsolatok.

Két vállalat (Insilico Medicine és Synthace) már elkezdte beépíteni a ChatGPT-t a rendszerükbe. A

gyógyszergyárakban való alkalmazás azonban még várat magára, ami az ipart jellemző szigorú szabályozásokkal magyarázható. A gyógyszergyártó vállalatok sokáig tartózkodtak digitális gyártási eljárások bevezetésétől, mivel úgy látták, hogy rendszereik, adataik és embereik nincsenek erre felkészülve. A nemzetközi gyógyszerügyi hatóságok ösztönzésére azonban az utóbbi években több gyógyszergyár is elkezdett kísérletezni a digitalizációval, azon belül a mesterséges intelligencia lehetséges alkalmazásaival a gyártás és minőség-ellenőrzés területén.

Összességében tehát a ChatGPT költségghatékony lehetőséget biztosít nagy mennyiségű adat kezelésére, elemzésére és új ismeretek generálására, ezáltal segítve a kutatóknak a gyógyszerkutatás és -fejlesztés felgyorsításában. Fontos azonban kiemelni, hogy a nyelvi modellek alkalmazhatóságát validálni kell más módszerekkel, mivel a modelleknek vannak korlátai.

A nyelvi modellek használata számos etikai szempontot is felvet, mint például hitelesség, plágium, adatvédelem és torzítások. Bizonyos esetekben megkérdőjelezhető az adatok minősége és megbízhatósága, akár hallucináció is felléphet. Emellett gondot jelenthet a jelenség tudományos háttérének korlátozott megértése is, mivel a nyelvi modell által generált válaszok nem mindig tükrözik a rendszerek valódi komplexitását.

A hibás válaszok előfordulása bizonyos mértékben javítható a modell finomhangolásával és részletesebb utasítások megadásával. Figyelembe véve ezeket a korlátokat és a modellek eredményeit, más kutatási eszközökkel együttesen alkalmazva a nagy nyelvi modellek rendkívül hasznos eszközei lehetnek a gyógyszeriparnak.

Köszönetnyilvánítás

A Doktoranduszi Kiválósági Ösztöndíj Program (DKÖP) által támogatott projekt a Kulturális és Innovációs Minisztérium Nemzeti Kutatási Fejlesztési és Innovációs Alapból nyújtott támogatásával, a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatallal kötött támogatási szerződés alapján valósult meg.

Rövidítésjegyzék

Rövidítés	Angol kifejezés	Magyar kifejezés
AEFI	<i>Adverse Event Following Immunisation</i>	Oltást követő nemkívánatos események
AEFI	<i>Adverse Event Following Immunisation</i>	Nemkívánatos események
AI	<i>Artificial Intelligence</i>	Mesterséges intelligencia
API	<i>Application Programming Interface</i>	Alkalmazásprogramozási felület
BERT	<i>Bidirectional Encoder Representations from Transformers</i>	
(BLSTM) RNN	<i>Bidirectional Long Short-Term Memory</i>	
EMA	<i>European Medicines Agency</i>	Európai Unióban az Európai Gyógyszerügynökség
FDA	<i>Food and Drug Administration</i>	Egyesült Államok Élelmiszer- és Gyógyszerügyi Hivatala
GMP	<i>Good Manufacturing Practice</i>	Helyes Gyógyszergyártási Gyakorlat
GPT	<i>Generative Pre-trained Transformer</i>	
LLM	<i>Large Language Model</i>	Nagy nyelvi modell
MedDRA	<i>Medical Dictionary for Regulatory Activities</i>	
MOF	<i>Metal Organic Frameworks</i>	Fémorganikus térhálók
NLP	<i>Natural Language Processing</i>	Természetes nyelvi feldolgozás
OGYÉI		Országos Gyógyszerészeti és Élelmezés-egészségügyi Intézet
PSUR	<i>Periodic Safety Update Reports</i>	Időközönként időszakos biztonságossági jelentés
QPPV	<i>Qualified Person Responsible for Pharmacovigilance</i>	Farmakovigilanciáért felelős személy
RLHF	<i>Reinforcement Learning with Human Feedback</i>	Visszajelzés alapú megerősítő tanulás
RNN	<i>Recurrent Neural Network</i>	Rekurrens neurális hálózat
UMLS	<i>Unified Medical Language System</i>	Egységes orvosi nyelvrendszer

WHO	World Health Organization	Egészségügyi Világszervezet
-----	---------------------------	-----------------------------

2. sz. táblázat: Rövidítésjegyzék
 Forrás: Saját szerkesztés

Irodalomjegyzék

- Abou-El-Enein, M.–Römhild, A.–Kaiser, D.–Beier, C.–Bauer, G.–Volk, H.-D.–Reinke, P. (2013) Good Manufacturing Practices (GMP) manufacturing of advanced therapy medicinal products: a novel tailored model for optimizing performance and estimating costs. *Cytotherapy*, Vol. 15. No. 3. pp. 362–383. doi: 10.1016/j.jcyt.2012.09.006
- Alkaissi, H.–McFarlane, S. I. (2023) Artificial Hallucinations in ChatGPT: Implications in Scientific Writing. *Cureus*, Vol. 15. No. 3. e35179. doi: 10.7759/cureus.35179
- Asselah, T.–Durantel, D.–Pasmant, E.–Lau, G.–Schinazi, R. F. (2021) COVID-19: Discovery, diagnostics and drug development. *Journal of Hepatology*, Vol. 71. No. 1. pp. 168–184. doi: 10.1016/j.jhep.2020.09.031
- Badini, S.–Regondi, S.–Frontoni, E.–Pugliese, R. (2023) Assessing the capabilities of ChatGPT to improve additive manufacturing troubleshooting. *Advanced Industrial and Engineering Polymer Research*, Vol. 6. No. 3. pp. 278–287. doi: 10.1016/j.aiepr.2023.03.003
- Biseda, B.–Mo, K. (2020) *Enhancing Pharmacovigilance with Drug Reviews and Social Media*. doi: 10.48550/arXiv.2004.08731
- Carpenter, K.A.–Altman, R.B. (2023) Using GPT-3 to Build a Lexicon of Drugs of Abuse Synonyms for Social Media Pharmacovigilance. *Biomolecules*, Vol. 13. No. 2. 387. doi: 10.3390/biom13020387
- Chakraborty, C.–Bhattacharya, M.–Lee, S. S. (2023) Artificial intelligence enabled ChatGPT and large language models in drug target discovery, drug discovery, and development. *Molecular Therapy - Nucleic Acids*, Vol. 33. pp. 866–868. doi: 10.1016/j.omtn.2023.08.009
- Cocos, A.–Fiks, A. G.–Masino, A. J. (2017) Deep learning for pharmacovigilance: Recurrent neural network architectures for labeling adverse drug reactions in Twitter posts. *Journal of the American Medical Informatics Association*, Vol. 24. No. 4. pp. 813–821. doi: 10.1093/jamia/ocw180
- Dreyer, N. A.–Blackburn, S.–Hliva, V.–Mt-Isa, S.–Richardson, J.–Jamry-Dziurla, A. et al. (2015) Balancing the Interests of Patient Data Protection and Medication Safety Monitoring in a Public-Private Partnership. *JMIR Medical Informatics*, Vol. 3. No. 2. doi: 10.2196/medinform.3937
- Fatani, B. (2023) ChatGPT for Future Medical and Dental Research. *Cureus*, Vol. 14. No. 4. e37285. doi: 10.7759/cureus.37285
- Fittler, A. (2012) Mit tehetnek a gyakorló gyógyszerészek a biztonságos gyógyszeralkalmazás érdekében? A farmakovigilancia és a gyógyszermellékhatás bejelentések szabályozásának változásairól. *Gyógyészerészet*, Vol. 56. No. 11. pp. 650–655.
- Golder, S.–Norman, G.–Loke, Y. K. (2015) Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. *British Journal of Clinical Pharmacology*, Vol. 80. No. 4. pp. 878–888. doi: 10.1111/bcp.12746
- Greenleaf, G. (2012) Global data privacy laws: 89 countries, and accelerating. In: *Privacy Laws & Business International Report*, Vol. 115. pp. 1–14.
- Gu, Y.–Tinn, R.–Cheng, H.–Lucas, M.–Usuyama, N.–Liu, X. et al. (2022) Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing. *ACM Transactions on Computing for Healthcare*, Vol. 3. No. 1. pp. 1–23. doi: 10.1145/3458754
- Hadi-Péter, E. S.–Hadi, K.–Hadi, F.–Szentirmai, B.–Hankó, B. (2010) Farmakovigilancia – mellékhatás bejelentés. *Gyógyészerészet*, Vol. 54. No. 5. pp. 259–264.
- Hazell, L.–Shakir, S. A. W. (2006) Under-Reporting of Adverse Drug Reactions. *Drug Safety*, Vol. 29. No. 5. pp. 385–396. doi: 10.2165/00002018-200629050-00003
- Heck, T. G. (2023) What artificial intelligence knows about 70 kDa heat shock proteins, and how we will face this ChatGPT era. *Cell Stress and Chaperones*, Vol. 28. No. 3. pp. 225–229. doi: 10.1007/s12192-023-01340-1
- Hole, G.–Hole, A. S.–McFalone-Shaw, I. (2021) Digitalization in pharmaceutical industry: What to focus on under the digital implementation process? *International Journal of Pharmaceutics: X*, Vol. 3. 100095. doi: 10.1016/j.ijpx.2021.100095
- Hussain, S.–Afzal, H.–Saeed, R.–Iltaf, N.–Umair, M. Y. (2021) Pharmacovigilance with Transformers: A

- Framework to Detect Adverse Drug Reactions Using BERT Fine-Tuned with FARM. *Computational and Mathematical Methods in Medicine*. Vol. 2021. 5589829. doi: 10.1155/2021/5589829
- Hussain, Z.–Sheikh, Z.–Tahir, A.–Dashtipour, K.–Gogate, M.–Sheikh, A.–Hussain, A. (2022) Artificial Intelligence–Enabled Social Media Analysis for Pharmacovigilance of COVID-19 Vaccinations in the United Kingdom: Observational Study. *JMIR Public Health and Surveillance*, Vol. 8. No. 5. e32543. doi: 10.2196/32543
- Juhi, A.–Pipil, N.–Santra, S.–Mondal, S.–Behera, J. K.–Mondal, H. (2023) The Capability of ChatGPT in Predicting and Explaining Common Drug–Drug Interactions. *Cureus*, Vol. 15. No. 3. e36272. doi: 10.7759/cureus.36272
- Kasneji, E.–Sessler, K.–Küchemann, S.–Bannert, M.–Dementieva, D.–Fischer, F., et al. (2023) ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, Vol. 103. 102274. doi: 10.1016/j.lindif.2023.102274
- Klein, A. Z.–Cai, H.–Weissenbacher, D.–Levine, L. D.–Gonzalez-Hernandez, G. (2020) A natural language processing pipeline to advance the use of Twitter data for digital epidemiology of adverse pregnancy outcomes. *Journal of Biomedical Informatics: X*, Vol. 112S. 100076. doi: 10.1016/j.jybinx.2020.100076
- Koromina, M.–Pandi, M. T.–Patrinos, G. P. (2019) Rethinking Drug Repositioning and Development with Artificial Intelligence, Machine Learning, and Omics. *OMICS A Journal of Integrative Biology*, Vol. 23. No. 11. pp. 539–548. doi: 10.1089/omi.2019.0151
- Kotschy, A. (2021) A gyógyszerkutatás új irányzatai: hatékonyság és biztonságosság. *Scientia et Securitas*, Vol. 2. No. 2. pp. 177–183. doi: 10.1556/112.2021.00028
- Krentzel, D.–Shorte, S. L.–Zimmer, C. (2023) Deep learning in image-based phenotypic drug discovery. *Trends in Cell Biology*, Vol. 33. No. 7. pp. 538–554. doi: 10.1016/j.tcb.2022.11.011
- Lamberti, M. J.–Wilkinson, M.–Donzanti, B. A.–Wohlhieter, G. E.–Parikh, S.–Wilkins, R. G.–Getz, K. (2019) A Study on the Application and Use of Artificial Intelligence to Support Drug Development. *Clinical Therapeutics*, Vol. 41. No. 8. pp. 1414–1426. doi: 10.1016/j.clinthera.2019.05.018
- Leaman, R.–Wojtulewicz, L.–Sullivan, R.–Skariah, A.–Yang, J.–Gonzalez, G. (2010) Extracting Adverse Drug Reactions from User Posts in Health-Related Social Networks. In: *Proceedings of the 2010 Workshop on Biomedical Natural Language Processing*, pp. 117–125. Uppsala, Sweden. Association for Computational Linguistics.
- Lee, S.–Woo, H.–Lee, C. C.–Kim, G.–Kim, J. Y.–Lee, S. (2023) Drug_SNSMiner: standard pharmacovigilance pipeline for detection of adverse drug reaction using SNS data. *Scientific Reports*, Vol. 13. No. 1. 3779. doi: 10.1038/s41598-023-28912-6
- Lengsavath, M.–Dal Pra, A.–de Ferran, A. M.–Brosch, S.–Härmark, L.–Newbould, V.–Goncalves, S. (2017) Social Media Monitoring and Adverse Drug Reaction Reporting in Pharmacovigilance: An Overview of the Regulatory Landscape. *Therapeutic Innovation and Regulatory Science*, Vol. 51. No. 1. pp. 125–131. doi: 10.1177/2168479016663264
- Liang, G.–Fan, W.–Luo, H.–Zhu, X. (2020) The emerging roles of artificial intelligence in cancer drug development and precision therapy. *Biomedicine and Pharmacotherapy*, Vol. 128. pp. 110255. doi: 10.1016/j.biopha.2020.110255
- Luo, Z.–Xie, Q.–Ananiadou, S. (2023) ChatGPT as a Factual Inconsistency Evaluator for Text Summarization. *Arxiv*, 2303.15621. doi: 10.48550/arXiv.2303.15621
- Magge, A.–Tutubalina, E.–Miftahutdinov, Z.–Alimova, I.–Dirkson, A.–Verberne, S. et al. (2021) DeepADEMiner: a deep learning pharmacovigilance pipeline for extraction and normalization of adverse drug event mentions on Twitter. *Journal of the American Medical Informatics Association*, Vol. 28. No. 10. pp. 2184–2192. doi: 10.1093/jamia/ocab114
- Mak, K. K.–Pichika, M. R. (2019) Artificial intelligence in drug development: present status and future prospects. *Drug Discovery Today*, Vol. 24. No. 3. pp. 773–780. doi: 10.1016/j.drudis.2018.11.014
- Mann, D. L. (2023) Artificial Intelligence Discusses the Role of Artificial Intelligence in Translational Medicine: A JACC: Basic to Translational Science Interview With ChatGPT. *JACC: Basic to Translational Science*, Vol. 8. No. 2. pp. 221–223. doi: 10.1016/j.jacbts.2023.01.001
- McCauley, S. M.–Christiansen, M. H. (2019) Language learning as language use: A cross-linguistic model of child language development. *Psychological Review*, Vol. 126. No. 1. pp. 1–51. doi: 10.1037/rev0000126
- Mohs, R. C.–Greig, N. H. (2017) Drug discovery and development: Role of basic biological research. *Alzheimer's and Dementia: Translational Research and Clinical Interventions*, Vol. 3. No. 4. pp. 651–657. doi: 10.1016/j.trci.2017.10.005
- Moon, S.–Chi, S.–Im, S. B. (2022) Automated detection of contractual risk clauses from construction specifications using bidirectional encoder representations from transformers (BERT).

- Automation in Construction*, Vol. 142. 104465. doi: 10.1016/j.autcon.2022.104465
- Pappa, D.–Stergioulas, L. K. (2019) Harnessing social media data for pharmacovigilance: a review of current state of the art, challenges and future directions. *International Journal of Data Science and Analytics*, Vol. 8. No. 2. pp. 113–135. doi: 10.1007/s41060-019-00175-3
- Park, S. H.–Hong, S. H. (2018) Identification of Primary Medication Concerns Regarding Thyroid Hormone Replacement Therapy From Online Patient Medication Reviews: Text Mining of Social Network Data. *Journal of Medical Internet Research*, Vol. 20. No. 10. pp. e11085. doi: 10.2196/11085
- Pimpalkhute, P. –Patki, A.–Nikfarjam, A.–Gonzalez G. (2014) Phonetic Spelling Filter for Keyword Selection in Drug Mention Mining from Social Media. *AMIA Joint Summits on Translational Science proceedings*, Vol. 2014. pp. 90–95.
- Powell, G. E.–Seifert, H. A.–Reblin, T.–Burstein, P. J.–Blowers, J.–Menius, J. A. et al. (2016) Social Media Listening for Routine Post-Marketing Safety Surveillance. *Drug Safety*, Vol. 39. No. 5. pp. 443–454. doi: 10.1007/s40264-015-0385-6
- Prakash Babu, Y.–Eswari, R. (2020) Identification of Medication Tweets Using Domain-specific Pre-trained Language Models. In: *Proceedings of the Fifth Social Media Mining for Health Applications Workshop*, pp. 128–130.
- Ray, P. P. (2023) ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, Vol. 3. pp. 121–154. doi: 10.1016/j.iotcps.2023.04.003
- Reynolds, L.–McDonell, K. (2021) Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–7. New York, NY, USA. doi: 10.1145/3411763.3451760
- Roche, V.–Robert, J. P.–Salam, H. (2023) A holistic AI-based approach for pharmacovigilance optimization from patients behavior on social media. *Artificial Intelligence in Medicine*, Vol. 144. 102638. doi: 10.1016/j.artmed.2023.102638
- Sarker, A.–Gonzalez, G. (2017) A corpus for mining drug-related knowledge from Twitter chatter: Language models and their utilities. *Data in Brief*, Vol. 10. pp. 122–131. doi: 10.1016/j.dib.2016.11.056
- Savage, N. (2023) Drug discovery companies are customizing ChatGPT: here's how. *Nature Biotechnology*, Vol. 41. pp. 585–586. doi: 10.1038/s41587-023-01788-7
- Stahl, B. C.–Eke, D. (2024) The ethics of ChatGPT – Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, Vol. 74. 102700. doi: 10.1016/j.ijinfomgt.2023.102700
- Stokel-Walker, C. (2023) ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, Vol. 613. No. 7945. pp. 620–621. doi: 10.1038/d41586-023-00107-z
- Sun, D.–Gao, W.–Hu, H.–Zhou, S. (2022) Why 90% of clinical drug development fails and how to improve it? *Acta Pharmaceutica Sinica B*, Vol. 12. No. 7. pp. 3049–3062. doi: 10.1016/j.apsb.2022.02.002
- Synthace. (2023) <https://www.synthace.com/news/synthace-platform-integrates-with-chatgpt> [Letöltve: 2023.09.28.].
- Tekumalla, R.–Banda, J. M. (2020) Characterizing drug mentions in COVID-19 Twitter Chatter. In: *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*, doi: 10.18653/v1/2020.nlpCOVID19-2.25
- Topaz, M.–Lai, K.–Dhopeshwarkar, N.–Seeger, D. L.–Sa'Adon, R.–Goss, F. et al. (2015) Clinicians' reports in electronic health records versus patients' concerns in social media: A pilot study of adverse drug reactions of aspirin and atorvastatin. *Drug Safety*, Vol. 39. No. 3. pp. 243–250. doi: 10.1007/s40264-015-0381-x
- Wang, H.–Ding, Y. J.–Luo, Y. (2023) Future of ChatGPT in Pharmacovigilance. *Drug Safety*, Vol. 46. No. 8. pp. 711–713. doi: 10.1007/s40264-023-01315-2
- Zhao, A.–Wu, Y. (2023) Future implications of ChatGPT in pharmaceutical industry: drug discovery and development. *Frontiers in Pharmacology*, Vol. 41. pp. 17–19. doi: 10.3389/fphar.2023.1194216
- Zheng, Z.–Zhang, O.–Borgs, C.–Chayes, J. T.–Yaghi, O. M. (2023) ChatGPT Chemistry Assistant for Text Mining and the Prediction of MOF Synthesis. *Journal of the American Chemical Society*, Vol. 145. No. 32. pp. 18048–18062. doi: 10.1021/jacs.3c05819
- Zhu, M.–Song, Y.–Jin, G.–Jiang, K. (2020) Identifying personal experience tweets of medication effects using pre-trained RoBERTa language model and its updating. In: *EMNLP 2020 - 11th International Workshop on Health Text Mining and Information Analysis, LOUHI 2020, Proceedings of the Workshop*, pp. 127–137. doi: 10.18653/v1/2020.louhi-1.14
- 15/2012. (VIII. 22.) EMMI rendelet az emberi alkalmazásra kerülő gyógyszerek farmakovigilanciájáról. 2005. évi XCV. törvény Az emberi alkalmazásra kerülő gyógyszerekről és egyéb, a gyógyszerpiacot szabályozó törvények módosításáról.

Szerzők

Dr. Bezerédi Imre PhD

rendőr alezredes
Nemzeti Közszerzőgálati Egyetem
Rendészettudományi Kar
tagozatparancsnok-helyettes

Erdélyi Dóra

Miskolci Egyetem
Állam-és Jogtudományi Kar
hallgató

Leiter Miklós Patrik

Miskolci Egyetem
Állam-és Jogtudományi Kar
hallgató

Papp Orsolya

rendőr alezredes
Miskolci Rendvédelmi Technikum
mesteroktató
Microsoft Innovatív Pedagógus Szakértő
Microsoft Innovatív Pedagógus Mesterképző
Microsoft Minősített Pedagógus

Péterfi Orsolya

Budapesti Műszaki és Gazdaságtudományi Egyetem
Vegyészmérnöki és Biomérnöki Kar
Szerves Kémia és Technológia Tanszék
PhD hallgató

Dr. Szilágyi Róbert PhD

Debreceni Egyetem
Gazdaságtudományi Kar
Alkalmazott Informatika és Logisztika Intézet
egyetemi docens

Dr. Végh János, az MTA doktora

Debreceni Egyetem
egyetemi tanár