ESZTERI DÁNIEL – PÉTERFALVI ATTILA

Amikor a gépeink tanulnak minket: avagy a mesterséges intelligencia alapú döntéshozatal és profilozás szabályozásának európai uniós törekvéseiről

Absztrakt:

Jelen tanulmány első része a gépi tanulás, tehát az önálló, automatizált döntések meghozatalára képes szoftverek adatalapú tanításának megfeleltethetőségét vizsgálja az Európai Unió 2018. május 25-től alkalmazandó általános adatvédelmi rendeletének (GDPR) egyes előírásai szempontjából. A témát az MI társadalmi hatásának általános bevezetésével indítjuk. Ezek után felvázoljuk a gépi tanulás néhány kulcsfogalmának alapvető technológiai hátterét és az ezzel összefüggő adatkezelés egyes adatvédelmi jogi szempontból releváns kérdéseit. Később bemutatásra kerülnek a GDPR vonatkozó előírásai, és azok alkalmazhatóságával kapcsolatos egyes kérdések és lehetséges megoldások. A tanulmány második részében bemutatjuk az adatalapú automatikus profilalkotás társadalmi hatásának, azaz a választói akarat és tudat indirekt befolyásolásának elhíresült példáját, az ún. Cambridge Analytica-botrányt, amely kapcsán a jelenség adatvédelmi jelentősége jól szemléltethető. Végül az utolsó fejezetben röviden bemutatjuk az EU új mesterséges intelligencia kódexének tervezetét és azt, hogy az új jogszabály hogyan próbálná meg szabályozni ezt az egyre több tudományos és szakmai vitát generáló jelenséget.

Kulcsszavak: automatizált döntéshozatal, profilalkotás, GDPR, mesterséges intelligencia, gépi tanulás, Cambridge Analytica, MI rendelet

Abstract:

The first part of the paper examines the compliance of data-driven machine learning, i.e., softwares capable of making autonomous, automated decisions with certain provisions of the European Union's General Data Protection Regulation (GDPR) applicable from 25 May 2018. We start the topic with a general introduction to the social impact of AI. Next, we outline the basic technological background and some key concepts of machine learning, then the legally relevant issues of such data processing. Later, the relevant provisions of the GDPR and some questions and possible solutions related to their applicability are presented. In the second part of the study, we present a famous example of the social impact of data-based automated profiling, thus the indirect influence of voters' will and consciousness in the so-called Cambridge Analytica scandal, in which the data protection significance of the phenomenon can be very well illustrated. In the final chapter, we briefly present the draft of the EU's new AI Code and how the new legislation would try to regulate this phenomenon generating more and more scientific and professional debate.

Keywords: automated decision making, profiling, GDPR, artificial intelligence, machine learning, Cambridge Analytica, AI Regulation

1. Bevezetés: Mesterséges intelligencia mint szabályozandó jelenség

A mesterséges intelligencia (MI) fejlesztése és üzemeltetése kapcsán az elmúlt időszakban elkezdődött végre az aktív tudományos diskurzus a jogtudomány művelői között is. Amíg öt-hat évvel ezelőtt még a téma felvetése is futurisztikusnak, illetve kissé idealistának hatott, ma már ott tartunk, hogy az Európai Unió hatályos adatvédelmi jogi szabályozása külön cikkekben próbálja meg szabályozni az automatizált (gépi) döntéshozatalt, sőt küszöbön van az MI alapú szoftverek fejlesztésére és üzemeltetésére vonatkozó általános uniós jogszabály elfogadása is.

Érdemes azonban azt is megjegyezni, hogy a jelenséggel foglalkozó napi hírekben, sőt akár tudományos konferenciák kérdésblokkjaiban is még mindig tetten érhető az a fajta ősi félelem, amelyet az autonóm, döntések meghozatalára képes szoftverek és gépek világától való idegenkedés jelent az ember számára.

Az elmúlt években eljutottunk odáig, hogy a jogalkotó nem halogathatta tovább a szabályozási kérdéseket: az MI olyan szintű társadalmi jelenséggé értett, amelynek szabályozásával foglalkozni kell. Ennek kiváló löketet adott a felhasználói profilozás és automatizált döntéshozatal témája körül kirobbant több ügy, pl. az ún. Cambridge Analytica-botrány, amellyel a cikkünkben részletesen is foglalkozni fogunk.

Egyelőre azonban ne szaladjunk ennyire előre. Nézzük először meg, hogy a gépek és adatok kapcsolata honnan indult és milyen hatással lehet az emberiségre, ahhoz pedig milyen tipikus társadalmi és jogi reakciók kapcsolódhatnak.

2. Miért félünk a gépektől? És miért nem kell félnünk?

Emberi lényként hajlamosak vagyunk arra, hogy a „gondolkodó gépet" egy ponton túl antropomorf, az élő szervezetekre jellemző tulajdonságokkal ruházzuk fel, végső soron pedig mint – felsőbbrendűnek hitt – új létformát az emberiségre veszélyt jelentő jelenségként azonosítsuk. A filozófiában ezt a jelenséget a hátborzongató völgy (*uncanny valley*) fogal-

mával írta le először az 1970-es években Masahiro Mori japán filozófus. Ezek szerint, ahogy egyre inkább emberszerűbbek lesznek a robotok, úgy nő velük szemben a rokonszenvünk – de egy ponton túl, amikor már nagyon emberszerűek, egyszer csak bizarrnak, hátborzongatónak és veszélyesnek látjuk őket.[1]

Az önálló tudatra ébredő és teremtőjét elpusztító mesterséges lény archetipikus képét felvázoló pesszimista irányzatok gyökerei a 20. század előtti szépirodalomban és folklórban is megtalálhatók (például Frankenstein története). Ráadásul az emberek és a mesterséges lények közötti konfliktus nem csak az irodalmi fikció szintjén jelent meg. Az ipari forradalom alatt az „emberek munkáját elvevő" gépektől való rettegés szülte például a géprombolók mozgalmát az 1810-es években.[2]

A legújabb pesszimista vagy *alarmista* elméletek alapját elsősorban az úgynevezett „technológiai szingularitás" problémája adja, amely Ray Kurzweil szerint *egy jövőbeli korszak, amelyben a technológiai változás üteme olyan gyors lesz, a hatása pedig olyan mély, hogy az emberi élet visszafordíthatatlanul átalakul.* Kurzweil szerint a szingularitás hatására megjelenő emberfeletti intelligencia pedig könnyen kiszoríthatja az embert a létezésből.[3]

Stuart Russel és Peter Norvig az MI-jelenséget elemző, összefoglaló munkájában bemutatott más, optimistább elméletek (például I. J. Good, vagy Moravec tanai) szerint az embereket leigázó MI víziója az ismeretlentől, ember- és természetfelettitől való ősi, alapvető rettegésből fakad, csakúgy, mint korábban a szellemektől vagy boszorkányoktól való félelem. Az optimisták szerint, ha az MI-t megfelelően, azaz olyan ágensekként tervezik, amelyek a gazdáik céljait teljesítik, akkor a jelenlegi tervezés lépésenkénti előrehaladásából származó MI-k szolgálni fognak, nem pedig leigázni.[4]

A két felfogást áthidaló *navigacionista* irányzat szerint a szingularitás mentén létrejövő intelligenciarobbanás eljövetele, ha nem is kerülhető el, de annak lefolyásában végső soron az emberiségnek lesz óriási szerepe és felelőssége. Ennek alapján a számítási, problémamegoldási képesség-

[1] Mori 2012, 99.

[2] Barthelmess–Furbach 2014, 5.

[3] Kurzweil 2014. Idézi: Marosán 2019.

[4] Russell–Norvig 2000, 26. fejezet.

ben az emberit meghaladó gépi intelligencia megfelelő irányba történő bölcs navigálása a jövő leglényegesebb kihívása. Az emberi felelősséget és tárgyilagosságot képviselő navigacionista álláspont az MI-fejlesztés felnőttkorát, a felelős szülő és tanító képét vetíti előre. A minden egyes technológiai fejlesztés mögött meghúzódó emberi felelősség fontosságát a téma jogi szempontú feldolgozása kapcsán sem lehet elégszer hangsúlyozni. A bölcs navigálás és fejlesztés az intelligens szoftverek adatalapú tanítása kapcsán érhető tetten a leginkább.[5]

3. Mesterséges intelligencia vs. mesterséges öntudat

A számítási képesség felgyorsulásával járó technológiai szingularitás eljövetele (bármilyen formában is kerüljön arra sor) ezért nem feltétlenül vetíti előre azt a fajta „gondolkodó gépet", amely majd közli az emberi megalkotójával, hogy evolúciós szempontból meghaladta őt.

A gép által végzett számítások kimenetének elsősorban mi emberek adunk és tulajdonítunk értelmet, a feltételezett jelentés komplexitása vagy autenticitása miatt látunk bele egyfajta tudatosnak vélt intelligenciát.

A jelenleg túlsúlyban lévő tudományos álláspont szerint az absztrakt gépi gondolkodásnak nem a mesterséges intelligencia, hanem a mesterséges tudatosság (*artificial consciousness* – AC) lenne az előfeltétele, amely viszont nem létezhet önazonosságra és önreflexióra képes ágens nélkül. Mindehhez a belső állapotra való szakadatlan referenciaképzés és az éppen aktuális külső állapot összevetése is szükséges. Platón ezt már több ezer évvel ezelőtt úgy fogalmazta meg, hogy a lélek folyamatosan önmagával folytatott hangtalan beszélgetése az, amit gondolkodásnak nevezünk.[6]

A tudatosság alapja az öntudat (vagy éntudat) megléte, amely alatt az önmagunkra való reflexiót, önmagunknak a környezettől való elkülönülését, elhatárolását és az ennek hatására kialakult énképet, illetve szemlé-

[5] Eszteri 2021, 189–190.

[6] Szathmáry–Miskolczi 2018, 44.

leti módot értjük. Az öntudat kialakulása, fejlődése, az én kiválása a környezet egységből lassú folyamat.

A csecsemőknek például még nincs éntudata. Jacques Lacan francia pszichoanalitikus például az ún. tükörstádiumhoz köti a csecsemők azon reakcióját, amikor a tükör előtt, saját tükörképükkel szembesülve önfelismerése utaló reakciókat adnak. Lacan szerint ez az „aha-élmény" és az éntudatra való rácsodálkozás eseménysora a csecsemőknél hat hónapos kortól következhet be.[7]

Ezzel szemben az intelligencia a problémák gyors és hatékony megoldására való képességet jelenti az információk észlelése, feldolgozása és az így megszerzett tudomás későbbi felhasználás céljából való elraktározása révén. Amíg az intelligencia modellezésében a szoftverek már rég hatékonyabbak és gyorsabbak, mint az emberek, mégsem képesek arra, hogy önmagukat, mint a környezetüktől elkülönült, elhatárolt entitást lássák (és nem csak azért, mert nincs a kezük ügyében egy tükör).

Z. Karvalics László kifejezését idézve a gépben nem zajlik információfeldolgozás, csak „kódmanipuláció". A gép jelműveleteket hajt végre, programjának megfelelően, de nincs a műveletre magára vonatkozó „metaszintje". Mint amikor valaki elsajátítja, hogyan kell összeadni, kivonni, szorozni, osztani, de nem tudja, miért, mikor, minek az érdekében van szükség erre. A szoftvernek tudati szinten nincsenek céljai, nincs akarata, nincsenek referenciapontjai, amelyekhez viszonyítva magára, a környezetre és a már korábban kialakult jelentésekre való tekintettel kellene új jelentéseket létrehoznia, és annak alapján döntést hoznia.[8]

A rendszerszemléletű információtudomány ismert művelője, Alva Noë az MI-t *pszeudo-intelligenciának* nevezi és így érzékelteti a különbséget az elő szervezetek és az MI között: „Egyetlen sejtnek élettörténete van; környezetté alakítja át azt a médiumot, amelyben találja magát, és ezt a környezetet értékes hellyé szervezi. Tápanyagot keres. Megcsinálja magát – és azzal, hogy megcsinálja magát, értelmet visz az univerzumba. A géppel ellentétben az amőbának van információja [önmagáról] – begyűjti és feldolgozza azt."[9]

[7] Lacan 1993, 5–6.
[8] Z. Karvalics 2015, 13.
[9] Idézi: Z. Karvalics 2015, 14.

Az MI által adott esetben egészen kiválóan modellezett problémamegoldó képesség összemosása az emberi önreflexív tudatossággal önmagának ellentmondó diskurzusokhoz vezet, amelyben az „alarmista", szingularitás fenyegető eljövetelét valló egyik fél és a szingularitást tagadó másik fél nagyon könnyen elbeszélhetnek egymás mellett.

Egyébként az alarmista szerzők közül is páran kiemelik, hogy az ellenséges MI létrejöttének megakadályozása érdekében az intelligens gépeket az emberi társadalom irányába „barátságos" entitásnak kell tervezni, ennek érdekében etikai elvek beépítését szorgalmazzák a programokba.[10] Itt azonban hangsúlyozni kell, hogy a szoftver valószínűleg ezekben az esetekben sem fog a tudatosság szintjén tisztában lenni a barátságosság, vagy empátia elvont fogalmainak valódi értelmével és jelentéstartalmával, a meghatározott célok keretei között általa végzett számítások eredményei egy emberi szemlélő számára azonban barátságosnak, az emberi fejlődést támogatónak fognak tűnni.[11]

Összefoglalva: az (ön)tudat és az intelligencia eltérő fogalmak, az MI kapcsán kialakult közbeszédben ezeket mégis hajlamosak vagyunk összemosni egymással.

De mégis hogyan kapcsolódik az MI jelensége a személyes adatok kezeléséhez? A következő pontokban ennek megvilágítására teszünk kísérletet.

4. A mesterséges intelligencia kapcsolata a személyes adatokkal

A jelen korban már nap mint nap használt MI alapú rendszerek, szoftverek és eszközök olyan új típusú megoldásokat nyújtanak, amelyek nagyon sok esetben a felhasználók személyes adatainak kezelésével, feldolgozásával járnak együtt. A lakossági felhasználásra szánt otthoni robotok, vagy az emberi viselkedést elemző okostelefonos applikációk folyamatosan monitorozzák a felhasználóik viselkedését és reakcióit annak érdekében,

[10] Goertzel–Pitt 2011, idézi: Pokol 2018, 55–56.
[11] Eszteri 2021, 191.

hogy minél tökéletesebben tudják kiszolgálni az igényeiket. Nem véletlen, hogy az ilyen modern technológiai megoldást használó eszközöknél és szolgáltatásoknál ma már szinte minden esetben fontos kulcsszó a személyre szabottság. A személyre szabottság mellett azonban egyre nagyobb az igény az olyan technológiákra is, amelyek képesek előre megjósolni a felhasználó igényeit. Ez már sokkal bonyolultabb döntési mechanizmusokat feltételez, amelyeket leginkább MI alapú öntanuló rendszerekkel lehet kiváltani.[12]

A norvég adatvédelmi hatóság (Datatilsynet) témába vágó jelentése úgy írja le az MI-t, mint egy olyan rendszert, amely képes a saját tapasztalatai alapján tanulni és a megszerzett tudást eltérő helyzetekben alkalmazni összetett problémák megoldására. A koncepció lényege, hogy az MI az általa „látott" (a gyakorlatban tulajdonképpen bele töltött) személyes adatokból tanul és dönt vagy „jósol".[13]

Az MI-t és az ún. gépi tanulást sokszor szinonimaként használják, holott a két jelenség eltérő fogalmakat takar. Az MI gyűjtőfogalomként szolgál, amely magában foglalja valamennyi olyan eljárást, amikor egy szoftver automatikusan hoz meg egy döntést. A gépi tanulás ehhez képest szűkebb fogalom, amely az MI fejlesztés egyik ágát jelenti. Ennek lényege, hogy a rendszer tapasztalatokból generál önálló tudást. A rendszer példa-adatokban, adatbázisokban keresett minták alapján képes önállóan, vagy emberi segítséggel szabályszerűségeket, szabályokat felismerni és meghatározni, majd az elsajátított tudásbázisban felfedezett szabályszerűségek alapján döntéseket hozni.[14]

5. Hogyan tanítjuk adatokkal a gépeket?

Az MI működésének adatvédelmi vizsgálata szempontjából talán az egyik legfontosabb terület ezért az úgynevezett „gépi tanulás" jelensége, amely során a szoftver a belé töltött adatok alapján „tanul" és hoz meg különböző döntéseket. A piacon ez legtöbbször úgy jelenik meg, hogy az alkalma-

[12] Ibid. 193.
[13] Datatilsynet 2018.
[14] Szepesvári 2005.

zott technológia gyakorlatilag képessé válik arra, hogy előre megjósolja az azt használó ember igényeit.

A gépi tanulás során az MI-rendszer által végzett adatkezelést három lépcsőre lehet bontani, amelyek a következők: [15]

a) Először nagy mennyiségű tesztadatot táplálnak be a rendszerbe, ebben az adathalmazban pedig az algoritmus megpróbál mintákat, hasonlóságokat keresni. Amennyiben az algoritmus talál ilyen azonosítható mintákat, úgy azokat megjegyzi és elmenti későbbi használat céljából. A megjegyzett és elmentett minták alapján ezek után a rendszer egy úgynevezett „modellt" generál. A rendszer a modell segítségével, a már azonosított minták alapján képes feldolgozni a később általa „látott" (a gyakorlatban betáplált, betöltött) éles adatokat.

b) Ezek után a rendszerbe újabb, „éles" adatokat töltenek fel, amelyek hasonlóak a tanuláshoz használt adatokhoz. A korábban generált modell alapján az MI eldönti, hogy az új adat mely megtanult mintázathoz hasonlít a leginkább.

c) A rendszer végül informál arról, hogy milyen döntést hozott az elsajátított mintázatok alapján a beletáplált új adatokkal kapcsolatban.

Fontos azt is megjegyezni, hogy a gépi tanulás során létrejövő modell nem feltétlenül tartalmazza a forrásadatokat, amelyek a tanulásának alapjául szolgáltak. A legtöbb esetben a tanulás alapjául szolgáló adatoktól függetlenül is tud működni a gépi tanulás során létrejött MI-rendszer.[16]

6. Az Európai Unió általános adatvédelmi rendeletének (GDPR) vonatkozó fogalmai

A gépi tanuláson alapuló rendszereket egyre gyakrabban használják személyes adatokon alapuló döntések meghozatalára. Az interneten megjelenő személyre szabott reklámok, hirdetések, és más tartalmak nagyon jó példák arra, hogy az emberi viselkedést elemző és abból tanuló algoritmusok hogyan működnek és hogyan használják a személyes adatainkat

[15] Datatilsynet 2018, 7.
[16] Ibid. 10.

a minél személyesebb, célzott tartalmak megjelenítésére. A profilalkotással szorosan összefügg az automatizált döntéshozatal fogalma is, mivel az adott személlyel kapcsolatos minél egyedibb profil algoritmikus úton létrejött döntések mentén alakul ki.

A GDPR nem határozza meg, hogy mit értünk mesterséges intelligencia vagy gépi tanulás alatt. A rendelet tartalmazza azonban több helyen is az *automatizált döntéshozatal* kifejezést, viszont annak fogalmát explicite nem határozza meg.

Az Európai Adatvédelmi Testület elődjének tekinthető ún. 29. cikk szerint működő Adatvédelmi Munkacsoport (angol rövidítéssel: WP29) vonatkozó iránymutatása szerint az automatizált döntéshozatal az a képesség, hogy „technológiai eszközök segítségével, emberi beavatkozás nélkül hoznak döntéseket."[17] A kizárólag automatizált döntéshozatalban tehát nincs emberi részvétel a döntési folyamatban. A gépi tanulás tulajdonképpen az automatizált döntéshozatal előszobájának tekinthető. E szerint a kizárólag automatizált, tehát a gép által meghozott döntést a legtöbb esetben az adatok valamilyen fajta automatikus értékelésének kell megelőznie. Ez az értékelés nagyon sok esetben a rendszer gépi tanulás során elsajátított és beazonosított mintázatai alapján történik.[18]

A GDPR-ban meghatározó elem továbbá a profilalkotás fogalma, amelyet – az automatizált döntéshozatallal ellentétben – már meghatároz a rendelet 4. cikk 4. pontja. A fogalom szerint a profilalkotás célja egy természetes személy személyes jellemzőinek értékelése. Általánosságban elmondható, hogy a profilalkotás egy természetes személyről (vagy természetes személyek csoportjáról) való információgyűjtést és a jellemzőik vagy a viselkedési mintáik értékelését jelenti annak érdekében, hogy bizonyos kategóriába vagy csoportba sorolja őt (vagy őket). A besorolás célja, hogy annak során elemezze az érintett érdeklődési körét, a tőle várható magatartást vagy bizonyos képességeit.[19] Később a felállított személyiségprofil alapján személyre szabott üzeneteket, vagy szolgáltatásokat lehet eljuttatni az érintettnek.

Fontos megjegyezni, hogy az automatizált döntéshozatal és a profilalkotás fogalmai nem azonosak teljes mértékben. Létezhet olyan automati-

[17] 29. cikk szerint működő Adatvédelmi Munkacsoport 2017, 8.
[18] Eszteri 2021, 199–200.
[19] 29. cikk szerint működő Adatvédelmi Munkacsoport 2017, 8.

zált döntéshozatali eljárás, amely nem minősül egyben profilalkotásnak, illetve profilalkotást is lehet végezni automatizált döntéshozatali mechanizmusok beépítése nélkül. A legtöbb esetben azonban a két fogalom kiegészíti egymást, így adatvédelmi szempontból indokolt az együttes tárgyalásuk.

7. Az automatizált döntéshozatal és profilalkotás szabályozása a GDPR-ban

A GDPR-ban az automatizált döntéshozatal és ezzel szoros összefüggésben a profilalkotás jelenségeivel kapcsolatosan a rendelet 22. cikke tartalmaz közös előírásokat. A cikk (1) bekezdése alapján az érintett jogosult arra, hogy ne terjedjen ki rá az olyan, *kizárólag* automatizált adatkezelésen – ideértve a profilalkotást is – *alapuló* döntés hatálya, amely rá nézve *joghatással* járna vagy őt *hasonlóképpen jelentős mértékben érintené.* Ez a rendelkezés a rendelet szövegezése ellenére valójában nem az érintett számára biztosított jogosultság, hanem az adatkezelőre vonatkozó általános tilalom, amely a kizárólag automatizált adatkezelésen alapuló döntéshozatali folyamatok alkalmazását tiltja. Ez a tilalom attól függetlenül fennáll, hogy az érintett tesz-e intézkedést a személyes adatainak kezelésével kapcsolatban. Főszabályként tehát a GDPR általános tilalmat állít fel a joghatással vagy hasonlóképpen jelentős hatással járó, kizárólag automatizált egyedi döntéshozatalra és profilozásra.[20]

A tisztán automatizált döntéshozatalra vonatkozó szabályokat továbbá csak azokban az esetekben kell alkalmazni, amikor az joghatással, vagy hasonló jelentős hatással jár az érintett természetes személyre nézve. A GDPR nem határozza meg a „joghatás" vagy a „hasonlóképpen jelentős hatás" fogalmakat, azonban a rendelet ezen megfogalmazása egyértelművé teszi, hogy a 22. cikk csak a súlyos következményt jelentő hatásokra terjed ki.[21]

20 Veale–Edwards 2018, 400.
21 Ibid. 401.

A joghatás megköveteli, hogy a gépi döntés befolyásolja valaki törvényes jogait. Joghatás lehet olyasmi is, ami befolyásolja a személy jogállását vagy szerződésen alapuló jogait. A WP29 szerint az ilyen jellegű hatásra való példák közé tartoznak a természetes személyekre vonatkozó azon automatizált döntések, amelyek eredményeként szerződést mondanak fel; a törvény által biztosított szociális ellátást – például gyermekkel kapcsolatos vagy lakhatási támogatást – ítélnek oda vagy tagadnak meg; megtagadják a belépést egy országba vagy megtagadják az állampolgárságot.[22]

Az automatizált döntéshozatal hatása az emberek törvényben vagy szerződésben lefektetett jogaira viszonylag egyértelműen körülhatárolható eseteket érint. Emellett azonban megjelenik a homályosabban megfogalmazott „hasonlóképpen jelentős hatás" fogalma is a GDPR 22. cikkében, mint szintén a tiltást megalapozó körülmény.

A GDPR (71) preambulumbekezdése tartalmaz ezzel a fogalommal kapcsolatban némi fogódzót, amely az alábbi példákat sorolja fel: „egy online hitelkérelem automatikus elutasítása" vagy „emberi beavatkozás nélkül folytatott online munkaerő-toborzás".

Nehéz pontosan meghatározni, hogy mit kell kellően *jelentős mértékűnek* tekinteni ahhoz, hogy elérje a küszöböt, azonban a WP29 szerint a következő döntések ebbe a kategóriába tartozhatnak: az egyén anyagi körülményeit befolyásoló döntések, például a hitelre való jogosultságát illetően; olyan döntések, amelyek befolyásolják az egyén egészségügyi szolgáltatásokhoz való hozzáférését; olyan döntések, amelyek megtagadnak valakitől egy foglalkoztatási lehetőséget, vagy valakit súlyos hátránynak tesznek ki; döntések, amelyek befolyásolják valakinek az oktatáshoz való hozzáférését, például egyetemi felvételét.[23]

A WP29 szerint az online fogyasztói profilok felépítésén alapuló célzott reklámra vonatkozó automatikus döntésnek legtöbbször nem lesz hasonlóan jelentős mértékű hatása a természetes személyekre (pl. ruházati hirdetések). Ebben a kategóriában is vannak azonban olyan adatkezelések, amelyek jelentős hatást gyakorolhatnak a társadalom bizonyos csoportjaira, például a kiszolgáltatott helyzetű felnőttekre.

22 29. cikk szerint működő Adatvédelmi Munkacsoport 2017, 22.

23 Ibid. 23.

Például, ha a felállított profil alapján egy személy valószínűsíthetően pénzügyi nehézségekkel küzd, és őt mégis rendszeresen magas kamatozású hitelekről szóló hirdetésekkel veszik célba, akkor potenciálisan további adósságot halmoz fel (feltéve, ha elfogadja az ilyen ajánlatokat).[24] Az ilyen esetekre már vonatkozik a 22. cikkben megfogalmazott általános tilalom. Fő szabály szerint így egy anyagi nehézségekkel küzdő fogyasztóról (gépi tanulással) alkotott profilt nem lehet azon célból felhasználni, hogy őt további anyagi kockázatvállalásra próbálják célzottan rávenni. Nem hivatkozhatnak arra az adatkezelést végző profilalkotók, hogy a hitel felvétele tőlük függetlenül az érintett döntése, mivel a profilalkotás – amelyen a fogyasztói döntés alapul – sem jogszerű.

A fent írtak szerint a 22. cikk (1) bekezdése általános tilalmat emel a joghatással vagy hasonlóképpen jelentős hatással járó kizárólag automatizált egyedi döntéshozatallal szemben. Léteznek azonban kivételek ezen általános tilalom alól, amelyeket a 22. cikk (2) bekezdése nevesít. Ezek szerint a tilalom nem alkalmazható abban az esetben, ha a döntés:

1., az érintett és az adatkezelő közötti *szerződés* megkötése vagy teljesítése érdekében szükséges;
2., meghozatalát az adatkezelőre alkalmazandó olyan *uniós vagy tagállami jog teszi lehetővé,* amely az érintett jogainak és szabadságainak, valamint jogos érdekeinek védelmét szolgáló megfelelő intézkedéseket is megállapít; vagy
3., az érintett kifejezett *hozzájárulásán* alapul.

Az első kivétel a szerződés teljesítése, amely alapján egy szerződés kapcsán létrejövő jogviszonyban az adatkezelők alkalmazhatnak automatizált döntéshozatali folyamatokat a szerződéssel összefüggő célokra. A WP29 szerint ebben az esetben az adatkezelőnek be kell tudnia mutatni, hogy az automatizált döntéshozatal alkalmazása a legmegfelelőbb adatkezelési módszer a szerződésben meghatározott célok eléréséhez. Ha a szerződéssel elérni kívánt célt – a tudomány és technológia állását és a megvalósítás költségeit is figyelembe véve – más módszerrel is el le-

[24] Ibid. 24.

het hatékonyan, a kockázatokkal arányosan érni, az már nem minősül szükségesnek és a beépített és alapértelmezett adatvédelem elvének[25] is ellentmond.

A második kivétel, ha az automatizált döntéshozatal lehetőségét az adott adatkezeléssel kapcsolatban az uniós vagy tagállami jog teszi lehetővé. A vonatkozó jogszabálynak az érintettek jogait és szabadságait, valamint jogos érdekeit védő megfelelő intézkedéseket is meg kell határoznia. A GDPR (71) preambulumbekezdése szerint ilyen lehet például, ha a jog a csalás és adóelkerülés megelőzése érdekében lehetővé teszi az állam számára, hogy automatizált döntéshozatali mechanizmusokat alkalmazzon.

Végül a harmadik kivétel, ha az automatizált döntéshozatal alkalmazása az érintett kifejezett hozzájárulásán alapul.[26]

Maga a GDPR nem határozza meg a „kifejezett hozzájárulás" fogalmát, ugyanakkor maga az „érintett hozzájárulása" fogalom[27] értelmében nyilatkozat vagy kifejező cselekedet szükséges a jogszerűséghez. Ezen felül a WP29 a hozzájárulással kapcsolatos iránymutatása útmutatást ad a hozzájárulás kifejezésének értelmezéséhez az alábbiak szerint.

A hozzájárulás kifejezett voltáról való meggyőződés legegyértelműbb módja a hozzájárulás írásbeli nyilatkozatban történő megerősítése. Az aláírt nyilatkozat azonban nem az egyetlen módja a kifejezett hozzájárulás megszerzésének. A WP29 szerint digitális vagy online kontextusban például előfordulhat, hogy az érintett elektronikus űrlap kitöltésével, elektronikus levél küldésével, az aláírását tartalmazó szkennelt dokumentum feltöltésével vagy elektronikus aláírás használatával is ki tudja állítani az előírt nyilatkozatot. Végül a hozzájárulás kétlépcsős ellenőrzésével is meg lehet győződni a kifejezett hozzájárulás érvényességéről (kétfaktoros autentikáció használata).[28]

[25] GDPR 25. cikk

[26] 29. cikk szerint működő Adatvédelmi Munkacsoport 2017, 25.

[27] GDPR 4. cikk 11. pont: „az érintett hozzájárulása": az érintett akaratának 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.

[28] 29. cikk szerint működő Adatvédelmi Munkacsoport 2018, 20–22.

8. Az automatizált döntéshozatal és profilalkotás társadalomra gyakorolt hatása, avagy a Cambridge Analytica-ügy

Az adatalapú gazdaság, profilalkotás és célzott hirdetések kapcsolatát jól bemutató egyik legismertebb ügy nyilvánosságra kerülésére 2018 elején került sor. Az ügy előzményei azonban még jóval a 2010-es évek elejére tehetők, amikor Aleksandr Kogan, a Cambridge University pszichológiai tanszékének kutatója, egy alkalmazást fejlesztett a Facebookra, amelynek a *„This is Your Digital Life"* (röviden: TIYDL) nevet adta. A TIYDL egy szórakoztatási célú személyiségelemző program volt, amely a felhasználójáról pszichológiai profilt készített. Az alkalmazás működtetését kutatási célból engedélyezte a Facebook mint a közösségi oldalra történő regisztráció során megadott személyes adatok adatkezelője. Ilyen applikációt egyébként bárki készíthet, a felhasználói adatokhoz való hozzáférés szabályait pedig a Facebook a mindenkor hatályos, az applikációk fejlesztői részére kiadott szabályzatában, a Facebook Platform Policy-ben határozza meg. Amennyiben az applikáció megfelel ezen szabályzat feltételeinek, úgy az elérhetővé válik a közösségi oldalon.[29]

Az alkalmazás használatának feltétele volt az egyes érintett felhasználók adatkezeléshez való hozzájárulása, továbbá, hogy megismerhessék az adataik felhasználásának céljait. A feltételek ismeretében mintegy 270 000 felhasználó használta a programot.[30] Utóbb azonban kiderült, hogy az alkalmazás nemcsak az azt használó érintettek adataihoz fért hozzá, hanem ismerőseikéhez is. A felhasználóról és ismerőseiről összeállított pszichológiai profil tartalmazta, hogy azok milyen politikai irányultságúak, milyen tartalmakat, vagy szereplőket követnek a Facebookon, mi a valláshoz való viszonyuk és hol helyezkednek el az ún. OCEAN-skálán,[31] amely egy betűszó öt tulajdonság angol nevéből.[32]

[29] Domokos 2021, 119–120.

[30] Németh 2021, 119.

[31] Az OCEAN skála ötdimenziós modelljének összetevői: *Openness* (érzelmekre, élményekre való nyitottság) – *Conscientiousness* (céltudatosság) – *Extraversion* (önbizalom, társaságkedvelés) – *Agreeableness* (kooperációra való hajlandóság) – *Neuroticism* (kellemetlen érzelmekre való hajlam). Lásd: http://medicalonline.hu/cikk/megelozheto_e_az_alzheimer_kor_

[32] Domokos 2021, 121.

Kogan harmadik személyeknek is átadta a teljes általa kezelt adatállományt, így a Cambridge Analytica, továbbá az Eunoia Technologies nevű cégeknek is. Ez ellentétben állt a Facebook Platform Policy-nek a TIYDL készítése idején hatályos változatával, amely megtiltotta az adatok harmadik személy részére való értékesítését a felhasználó hozzájárulása nélkül, továbbá az ismerősök adatainak felhasználását „a készítő saját céljaira".[33]

Azt, hogy a program az érintett felhasználók ismerőseinek adatait is gyűjti és kezeli, továbbá, hogy azokat Kogan harmadik feleknek is átadta 2015-ben észlelte a Facebook és ezért eltávolította az alkalmazást az oldalról. Egyúttal felbontották a szerződésüket Kogannel, és írásbeli igazolást kértek tőle, valamint az adattovábbítások címzettjeitől, hogy megsemmisítették az általuk jogellenesen kezelt személyes adatok teljes állományát. Az érintett gazdasági társaságok a kért nyilatkozatokat állítólag be is nyújtották a Facebook részére,[34] magát a törlési folyamatot azonban a Facebook nem ellenőrizte.[35]

Ezt követően érkezett el 2018 márciusa, amikor a sajtóban nyilvánosságra kerültek azon tényfeltáró cikkek, amelyekben Christopher Wylie, a sajtónak nyilatkozó korábbi munkavállaló azt állította, hogy a Cambridge Analytica nem csak, hogy nem semmisítette meg az általa jogellenesen kezelt személyes adatállományt, de azokat felhasználva a 2016-os amerikai elnökválasztási kampányban aktívan célzott bizonyos befolyásolható választói csoportokat politikai hirdetésekkel a korábban felállított pszichológiai profiljuk alapján. Ennek eredményeképpen sikeresen befolyásolták ezeket a választókat az ún. „billegő" választási körzetekben abba az irányba, hogy inkább Donald Trump republikánus elnökjelölt kampányát támogassák az elnökválasztáson és rá szavazzanak. Ugyanebben az évben szintén hatást gyakorolt a cég hasonló hirdetéseken keresztüli pszichés befolyásolási módszerekkel az Egyesült Királyság EU-tagságáról szóló népszavazására (Brexit-népszavazás). Wylie szerint a Cambridge Analytica ezen kampányok során mintegy 87 millió Facebook felhasználói adatait kezelte.[36]

[33] Facebook Platform Policy, II. 4. pont. Lásd: https://bit.ly/3rioTYH

[34] Németh 2021, 119.

[35] Domokos 2021, 123.

[36] Lásd: https://www.cnbc.com/2018/04/10/facebook-cambridge-analytica-a-timeline-of-the-data-hijacking-scandal.html

Az ügy következményeként 2018. október 24-én az Egyesült Királyság adatvédelmi hatósága (*Information Commissioner's Office,* röviden: ICO) a jogsértés elkövetésekor hatályos adatvédelmi jogszabályok szerinti meghatározott maximum bírságot (500 000 GBP) szabta ki[37] a Facebookra a Cambridge Analytica-botrány kapcsán.[38] 2019. júliusában pedig az USA versenyhatósága (*Federal Trade Commission,* röviden: FTC) 5 milliárd USD bírságot szabott ki a Facebookra a Cambridge Analytica-ügy nyomán indult vizsgálat eredményeképp.[39]

9. A Cambridge Analytica-ügy rövid értékelése az algoritmusok átláthatósága szempontjából

A fenti botrány adatvédelmi jogi szempontú értékelése kapcsán először is szükséges kitérni az úgynevezett mikrotargetálás fogalmára. Ennek lényege, hogy a begyűjtött személyes adatok (pl. böngészési szokások, megtekintett vagy lájkolt tartalmak, közösségi oldalon folytatott kommunikáció) alapján felállított felhasználói profil segítségével azonosítható egy specifikus célcsoport vagy személy érdeklődési köre és ennek alapján személyre szabott üzenet/tartalom küldhető számára az interneten keresztül.

A mikrotargetáláshoz szükséges adatelemzés és profilalkotás szinte teljes mértékben automatikus módon történik. Ennek során az automatizált döntéshozó algoritmus az érintettről begyűjtött adatok alapján meghatározza, hogy pontosan milyen típusú tartalmat, milyen gyakorisággal érdemes a részére küldeni, befolyásolva ezzel fogyasztói vagy politikai döntéseit, szokásait. Ha például valaki egy bizonyos termékcsaládot lájkol, egy zenei stílushoz tartozó előadókat hallgat vagy politikai ideológiát valló közszereplők tevékenységét követi, akkor számára az algoritmus ilyen és ehhez hasonló tartalmakat fog a továbbiakban megjeleníteni.[40]

[37] Lásd: https://ico.org.uk/media/action-weve-taken/mpns/2260051/r-facebook-mpn-20181024.pdf

[38] Lásd: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2018/10/facebook-issued-with-maximum-500-000-fine/

[39] Lásd: https://www.bbc.com/news/world-us-canada-48972327

[40] Domokos 2021, 122.

A mikrotargetálás módszere szorosan összefügg a Richard Thaler nevéhez fűződő ún. döntéstervezés- (*choice architecture*) és befolyásolás-elmélettel (*nudge theory*). Az angol *nudge* kifejezésnek a magyar nyelvben leginkább a noszogatás szó feleltethető meg, lényege a döntés előtt álló egyén döntésének indirekt módszerekkel valamilyen irányba való terelése, ösztönzése. Thaler szerint a *nudge* sosem jelenthet manipulációt, mindössze „enyhe orientációt".[41]

A Cambridge Analytica-ügy kapcsán a fenti fogalmaknak azért van különös jelentősége, mivel a TIYDL Facebookos felhasználóiról (és ismerőseikről) összeállított személyiségprofilokat pontosan ilyen automatikus mikrotargetált politikai hirdetések eljuttatására használták fel a későbbiekben, hogy az érintettek döntéseinek bizonyos irányba való terelgetésével indirekt módon befolyásolják a választásokat.

A Cambridge Analytica általi jogosulatlan adatkezelések pontos intervalluma a rendelkezésre álló információk alapján nem határozható meg, de bizonyosan a GDPR 2018. május 25-i alkalmazandóvá válása előtt történtek: a személyes adatok gyűjtése 2010–2015 közé, míg a választók befolyásolása célzott tartalmakkal a 2016-os évre tehető. Az adatkezeléshez szükséges jogalap igazolása, a transzparens, átlátható adatkezelés követelménye és az adatkezelésről való előzetes tájékoztatás megléte azonban a GDPR alkalmazása előtt is követelmények voltak mind a korábban hatályos magyar adatvédelmi jogszabály,[42] mind a nemzetközi adatvédelmi sztenderdek alapján. Az átláthatóság követelményét a GDPR is kifejezetten nevesíti az 5. cikk (1) bekezdés a) pontjában. Ezek szerint a személyes adatok kezelését jogszerűen és tisztességesen, valamint az *érintett számára átlátható* módon kell végezni.

A felhasználók saját döntésük alapján vették igénybe az applikációt, a használat előtt pedig hozzájárultak ahhoz, hogy az alkalmazás és annak „üzemeltetője" hozzáférjen az általuk a regisztráció során megadott személyes adatokhoz. A felhasználók adatainak kezelése szempontjából tehát az alkalmazás működése még jogszerűnek is volt tekinthető. Az azonban, hogy az alkalmazást használó felhasználók körén kívüli további érintetti kör (a felhasználók ismerősei) személyes adataihoz is hoz-

[41] Deli–Kocsis–Muhari 2021, 237–238.

[42] Az információs önrendelkezési jogról és az információszabadságról szóló 2011. évi CXII. törvény a GDPR alkalmazása előtt hatályos változata.

záférése volt az alkalmazás üzemeltetőjének már nem lehetett jogszerű. Ennek oka, hogy a felhasználók ismerősei már nem járultak hozzá adataik ilyen célú kezeléséhez és erről semmilyen tájékoztatást nem kaptak. Az adatkezelés átláthatósága és jogszerűsége így ezen mozzanat kapcsán sérült. További jogsértő adatkezelési körülmény, hogy sem a felhasználók, sem ismerőseik nem voltak tisztában adataik harmadik felekkel való megosztásával kapcsolatban.

Mind a TIYDL-ot fejlesztő szolgáltató, mind a Facebook által elkövetett mulasztások kapcsán kiemelhető a transzparencia hiánya, amely hosszú évek óta a közösségi oldal működésével szembeni kritikák egyik központi eleme: az adatkezelés belső szabályai és informatikai keretei éppúgy ismeretlenek a közvélemény és a jogalkalmazók számára, mint például a célzott hirdetéseket kiosztó algoritmusok működése.[43] Ezt a problémát egyébként maga a GDPR is igyekszik rendezni, amikor különböző átláthatósági és a tájékoztatással kapcsolatos különös követelményeket állapít meg az automatizált döntéshozatal és profilalkotás kapcsán.

Ezek szerint a GDPR tájékoztatási kötelezettséget ír elő az adatkezelő részére a kizárólag automatizált adatkezelésen alapuló, joghatással vagy hasonlóan jelentős hatással járó döntéshozatallal kapcsolatban, amelyet személyes adatok kezelésével hoznak. A rendelet beleérti ebbe a körbe az ilyen adatkezelésen alapuló profilalkotást is.[44] Ennek keretében a következő három információt kell közölni az érintettel:

1) tájékoztatni kell az ilyen típusú személyes adatkezelés tényéről;
2) érdemi tájékoztatást kell adni az alkalmazott logikáról;
3) és végül arról is, hogy az adatkezelés milyen jelentőséggel és milyen várható következményekkel bír az érintettre nézve.[45]

Az automatizált egyedi döntéshozatal tényének közlése viszonylag egyszerű követelmény, ennek keretében elég, ha az adatkezelő arról tájékoztat, hogy ilyen típusú adatkezelésre kerül sor. Fontos, hogy az érintett arról is tudomással bírjon, ha az automatizált egyedi döntéshozatal kapcsán egyben profilalkotásra is sor kerül.

Az alkalmazott logikáról való tájékoztatás miként je már több kérdést vet fel. Ez főleg az előző pontokban bemutatott gépi tanulási módszerek

[43] Klein–Tóth (szerk.) 2018, 50.
[44] GDPR 15. cikk (1) bekezdés h) pont.
[45] GDPR 13. cikk (2) bekezdés f) pont.

esetében jelenthet nagy kihívást az adatkezelő részére, mivel az sokszor rendkívül összetett, nagyon nehezen átlátható adatkezelési folyamatokon alapul.

A GDPR szerint az adatkezelőknek „érdemi információt" kell adniuk az alkalmazott logikáról, ráadásul azt világosan és közérthetően megfogalmazva kell nyújtaniuk. Önmagában így például nem lehet elég az, ha az adatkezelő csak általánosságban közli, hogy pl. neurális hálózaton alapuló rendszert üzemeltet, mivel az érintett így érdemben vajmi keveset fog felfogni abból, hogy mi történik az adatkezelés során a személyes adataival.[46]

Az érdemi információ viszont azt sem jelenti, hogy feltétlenül bonyolult magyarázatot kell nyújtania az alkalmazott algoritmusokról, sem azt, hogy az algoritmust teljes egészében fel kellene tárnia az adatkezelőnek. A technológia részletes bemutatása ugyanis a legtöbb esetben lerontaná a tájékoztatás közérthetőségét és hátráltatná a befogadását.[47] Emellett maga a GDPR is kimondja, hogy az alkalmazott logikáról való tájékoztatás nem érinti az üzleti titkokat vagy a szellemi tulajdont, így a szoftverek védelmét biztosító szerzői jogokat.[48] A technológia komplexitása természetesen nem lehet mentség a tájékoztatás teljes mellőzésére sem.

A GDPR ezen előírásai megkövetelik az automatizált döntéshozatalon alapuló profilalkotást használó adatkezelőktől, így többek között a mikrotargetálást használó hirdetéseket megjelenítő oldalaktól is, hogy átlátható tájékoztatást nyújtsanak az ilyen típusú adatkezelésről. Az ennek való megfelelés immár a GDPR jogi rezsimje alatt kulcskérdés, így a jogalkotó is felismerte az ilyen módszerek magánszférára jelentett nagy hatását. Remélhetőleg az előírások nem csak „kirakatjogszabálynak" lesznek jók, hanem ténylegesen be is fogják azokat tartani az adatkezelők a profilalkotás során.

Az ilyen és ehhez hasonló rendszerek fejlesztése és üzemeltetése azonban nem csupán adatvédelmi, adatkezelési kérdés, amelyet immár az Európai Uniós jogalkotó is felismert. A mesterséges intelligencia alapú rendszerek szabályozását vetíti előre az Európai Bizottság 2021. április 21-én nyilvánosságra hozott rendelettervezete, amelyet a következő pontban mutatunk be röviden.

[46] Eszteri 2019, 679–680.

[47] Péterfalvi–Révész–Buzás (szerk.) 2018, 158.

[48] GDPR (63) preambulumbekezdés.

10. Az új mesterséges intelligencia rendelettervezet

Az Európai Bizottság által nyilvánosságra hozott rendelettervezet, amely a GDPR-hoz hasonlóan közvetlenül alkalmazandó jogszabály lenne, valamennyi EU-s tagállamban egységesen végrehajtandó jogszabályként szabályozná a mesterséges intelligencia fejlesztést.

A nyilvánosságra hozott tervezet célja a Bizottság sajtóközleménye szerint, hogy Európa a megbízható MI globális központjává váljon. A tervezet az MI-ként való besorolásra három feltétel együttes teljesülését írja elő. Először is az MI-nek meghatározott technológiákat kell alkalmaznia, másodszor az ember által kijelölt célokat önállóan kell tudnia követni, majd végül olyan kimeneteleket kell tudnia produkálni, amelyekkel „befolyásolja" a környezetet. Zódi Zsolt idevágó írása alapján az utóbbi két kritérium tulajdonképpen az „autonómia" pontos meghatározásának kísérlete. Az új kódex tervezete egyébként a gépi tanuláson alapuló rendszereken kívül még két másik technológiacsoportot is megjelöl, amelyekre kiterjed a hatálya. Ezek a tudásreprezentáción alapuló és a statisztikai rendszerek. Zódi szerint ennek oka az lehet, hogy ily módokon is építhetőek olyan rendszerek, amelyek kimenetei a bonyolultságuk, összetettségük és a feldolgozott adatok mennyisége miatt nem tűnnek determinisztikusnak.[49]

A kódex ezen felül kockázatalapú megközelítést alkalmaz az MI-k besorolása szempontjából, amely összesen négy nagy kategóriába igyekszik felosztani a rendszereket:

a) Az első kockázati kategória, amely az elfogadhatatlanul magas kockázatúként besorolt rendszereket tartalmazza. Ezek alatt olyan MI-ket ért, amelyek egyértelműen veszélyeztetik az emberek biztonságát, megélhetését és jogait. Ide érti például a szabályozási koncepció az olyan rendszereket vagy alkalmazásokat, amelyek a felhasználók „szabad akaratának megkerülése érdekében" manipulálják az emberi viselkedést, valamint ide tartoznak az olyan rendszerek is, amelyek lehetővé teszik a kormányok általi „társadalmi pontozást".[50]

[49] Zódi 2021.

[50] Európai Bizottság 2021.

Az előbbi kategóriába sorolható az előző pontokban bemutatott Cambridge Analytica által végzett profilozás és mikrotargetálás alapú adatkezelés is, hiszen a felhasználók itt nem voltak azzal tisztában, hogy adataik által felállított profilokon keresztül igyekeznek őket befolyásolni politikai szempontból. Az utóbbit, mint tiltást igénylő kategóriát pedig minden bizonnyal a Kínai Népköztársaságban kifejlesztett és tesztelt ún. társadalmi kreditrendszer ihlette.[51]

b) A második, azaz magas kockázati kategóriába sorolja be a kódextervezet azon MI-technológiákat, amelyeket összesen kilenc olyan területen és/vagy célból alkalmaznak, amelyek magas kockázatot jelentenek az emberek egyes alapvető jogaira nézve. Ezek a területek:

- kritikus infrastruktúrák (pl. közlekedés),
- oktatás vagy szakképzés (pl. vizsgák pontozása),
- egyes termékek biztonsági berendezései (pl. robotsebészet),
- foglalkoztatás és munkavállalók irányítása (pl. munkaerő-felvételhez önéletrajz válogatás),
- alapvető magán- és közszolgáltatások (pl. hitelbesorolás),
- bűnüldözés (pl. bizonyítékok megbízhatóságának értékelése),
- menekültügy és határellenőrzés (pl. úti okmány valódiságának ellenőrzése),
- igazságszolgáltatás és demokratikus folyamatok (pl. törvények konkrét tényállásra való alkalmazása),
- végül az összes távoli biometrikus azonosító rendszer magas kockázatúnak minősül a kódextervezet szerint. Használatuk bűnüldözési célból nyilvános helyeken, valós időben fő szabály szerint tilos lenne. Ezen tilalom alól néhány kivételes esetben enged csak eltérést a szabályozási koncepció (pl. eltűnt gyermek felkutatása, közvetlen terrorveszély vagy súlyos bűncselekmény megelőzése) és azt is bírói vagy más független hatósági engedélyhez köti.[52]

A fenti kategóriákba besorolható MI-rendszereknek forgalomba hozataluk előtt szigorú kötelezettségeknek kell megfelelniük. A tervezet előírja, hogy valamennyi ilyen rendszernek a fejlesztése során megfelelő kockázatértékelési és csökkentési folyamatokon kell keresztülmennie.

[51] Kollár 2020.

[52] Európai Bizottság 2021.

Az MI fejlesztése során használt adatkészleteknek kiváló minőségűeknek kell lenniük és az eredmények nyomon követhetősége miatt minden tevékenységet naplózni kell, továbbá részletes dokumentációnak kell rendelkezésre állnia a megfelelőség értékelésével kapcsolatban. A tervezet előírja továbbá a felhasználók egyértelmű és érthető tájékoztatását, az emberi felügyelet szükségességét, valamint egyfajta alapelvi éllel a rendszer megbízhatóságát, pontosságát és biztonságos működésének követelményét.[53]

c) A tervezet korlátozott kockázatú MI-nek sorolja be az olyan rendszereket, amelyek használata során a felhasználóknak tisztában kell lenniük azzal, hogy nem emberrel, hanem egy géppel kommunikálnak (pl. csevegőrobotok). Ezen átláthatósági kötelezettségre valószínűsíthetően azért van szükség, hogy a felhasználókat ne tévessze meg a program és tisztában legyenek azzal, hogy nem egy másik ember van a „monitor túloldalán."

d) A tervezet végül a minimális kockázati kategóriába sorolja be az MI-k legnagyobb részét kitevő olyan rendszereket, amelyek használata a felhasználók jogaira és biztonságára nézve szinte alig hordoz kockázatokat. A kódex szabadon lehetővé teszi ezen rendszerek használatát és nem tartalmaz rájuk nézve beavatkozó intézkedéseket, így azokat gyakorlatilag kivonja a hatálya alól. Ilyen MI-re példa a spamszűrők vagy videójátékok használata.

A kódex kapcsán nemrég az Európai Adatvédelmi Testület (EDPB) és az Európai Adatvédelmi Biztos (EDPS) is kifejtette közös véleményét, amely általánosságban szintén üdvözli a tervezetet. Az EDPB néhány területen – így például a távoli biometrikus azonosítás terén – azonban szigorítaná a szabályokat. A vélemény fő szabály szerint megtiltaná például az olyan távoli biometrikus azonosító rendszerek használatát, amelyek alkalmasak arra, hogy az érintetteket tulajdonságaik alapján kategóriákba sorolja származás, nem, szexuális orientáció alapján, mivel ez könnyen vezethet hátrányos megkülönböztetéshez.[54]

A fentieken túl az Európai Tanács is üdvözölte a területre vonatkozó szabályozást. A lengyel és a cseh szenátus azonban szintén aggodalmukat fejezték ki a biometrikus azonosítási rendszerek közterületeken történő,

[53] Ibid.

[54] Európai Adatvédelmi Testület – Európai Adatvédelmi Biztos 2021.

a tervezet értelmében engedélyezett használatával kapcsolatban, így az EDPB álláspontjához hasonlóan a jelenleginél szigorúbb megközelítést szorgalmaznak a Bizottságnak írt levelükben.[55]

A kódextervezet az elkövetkezendő időszakban minden bizonnyal igen sűrűn fogja alapját képezni további szakmai és tudományos diskurzusoknak, amíg az teljesen elfogadásra nem kerül az Európai Unió által. Általánosságban elmondható, hogy a kockázatalapú megközelítés, és a tiltott, valamint magas kockázatú besorolásra tartozó rendszerek viszonylag szűk köre előremutató és kellően rugalmas szabályozást sejtet.

11. Összegzés

A mesterséges intelligencia fejlesztés és ezen belül is a gépek adatalapú tanítása révén láthattuk, hogy a szoftver által hozott döntések és viselkedés a tanításhoz felhasznált adatkészleteken múlik. A szoftverfejlesztő és a rendszer üzemeltetőjének felelőssége ezért az ilyen rendszerek kapcsán óriási. A terület pedig minden bizonnyal az EU új MI rendelettervezete miatt pedig csak még hangsúlyosabb lesz a jövőben.

Az automatizált döntéshozatal és a profilalkotás során nagyon fontos követelmény, hogy a tanításhoz használt adatkészletek megfelelő minőségűek legyenek, amelyet az adatbázisok gondos előzetes kiválogatásával, az adatok megfelelő címkézéssel lehet elérni. Alapvetően téves feltevés tehát, hogy minél több adattal dolgozik a gépi tanuló algoritmus, annál hatékonyabban fog tudni működni és dönteni a későbbiekben.[56] Az adatkészletek gondos előzetes kiválogatása és a szükséges mértékűre való szűkítése jellemzően hatékonyabb döntéshozó rendszereket eredményez, amelyet az eddigi tudományos álláspontok is megerősítenek, valamint az EU új rendelettervezete is elvi éllel mondja ki, mint alapvető követelmény. A kevesebb sokszor ezért itt is több, mint ahogy szokták mondani.

A fentieken túl a rendszerek éles működése során kezelt személyes adatokkal kapcsolatban alapvető követelmény a megfelelő adatkezelé-

[55] Pethő 2021.

[56] Datatilsynet 2018, 11.

si jogalap igazolása, az adatminimalizálás elvének figyelembevétele és a rendszer átlátható, megismerhető működésének követelménye, amely során az alkalmazott logikáról való tájékoztatás az egyik legfontosabb kulcselem.

Összességében elmondható, hogy az MI és adatalapú az automatizált döntéshozatal kérdésköre általános jogi szabályozási szempontból még mindig csak bontogatja a szárnyait, azonban az elmúlt időszakban megszületett konkrét szabályozási koncepciók előremutatónak tűnnek. A szabályozás gyakorlatban való alkalmazhatósága és hatékonysága az elkövetkezendő évek kulcskérdése lesz. A magunk részéről kíváncsian várjuk a joggyakorlati fejleményeket.

Irodalom

29. cikk szerint működő Adatvédelmi Munkacsoport (WP29) 2017: *Iránymutatás az automatizált döntéshozatallal és a profilalkotással kapcsolatban a 2016/679 rendelet alkalmazásához.* 2017. https://naih.hu/files/wp251rev01_hu.pdf

29. cikk szerint működő Adatvédelmi Munkacsoport (WP29) 2018: *Iránymutatás az (EU) 2016/679 rendelet szerinti hozzájárulásról (WP259rev.01.).* 2018. 04. 10. http://naih.hu/files/wp259-rev-0_1_HU.PDF

Barthelmess, Ulrike – Furbach, Ulrich 2014: Do We Need Asimov's Laws? In: *Lecture Notes in Informatics.* Bonn, Gesellschaft für Informatik.

Datatilsynet 2018: *Artificial intelligence and privacy.* Report, January 2018. https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf

Deli Gergely – Kocsis Réka – Muhari Nóra 2021: Akarva-akaratlanul – az adatvédelem és az akaratszabadság dilemmái. In: Török Bernát – Ződi Zsolt (szerk.) 2021: *A mesterséges intelligencia szabályozási kihívásai – Tanulmányok a mesterséges intelligencia és a jog határterületeiről.* Ludovika Egyetemi Kiadó, Budapest.

Domokos Márton 2021: Globális törésvonalak – a Cambridge Analytica-ügy. In: Szabó Endre Győző (szerk.): *Az Infotörvénytől a GDPR-ig.* Ludovika Egyetemi Kiadó, Budapest.

Eszteri Dániel 2021: A gépek adatalapú tanításának megfeleltetése a GDPR egyes előírásainak. In: Török Bernát – Ződi Zsolt (szerk.) 2021: *A mesterséges intelligencia szabályozási kihívásai – Tanulmányok a mesterséges intelligencia és a jog határterületeiről.* Ludovika Egyetemi Kiadó, Budapest.

Eszteri Dániel 2019: Hogyan tanítsuk jogszerűen a mesterséges intelligenciánkat. In: *Magyar Jog 2019/12.*

Európai Adatvédelmi Testület – Európai Adatvédelmi Biztos 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/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf

Európai Bizottság 2021: *A digitális korra felkészült Európa: A Bizottság új szabályokat és intézkedéseket javasol a kiválóságra és bizalomra épülő mesterséges intelligencia terén.* https://ec.europa.eu/commission/presscorner/detail/hu/IP_21_1682

Goertzel, Ben – Pitt, Joel 2011: Nine Ways to Bias Open-Source AGI Toward Friendliness. In: *Journal of Evolution and Technology* 22. (2011) 1. 116–131.

Klein Tamás – Tóth András (szerk.) 2018: *Technológia jog – Robotjog – Cyberjog.* Budapest, Wolters-Kluwer Kft.

Kollár Csaba 2020: Kína és a társadalmi kreditrendszere. In: *Hadtudomány* 2020/2. https://www.mhtt.eu/hadtudomany/2020/2020_2szam/079-097_Kollar.pdf

Kurzweil, Ray 2014: *A szingularitás küszöbén.* Budapest, Ad Astra.

Lacan, Jacques 1993: A tükör-stádium mint az én funkciójának kialakítója, ahogyan ezt a pszichoanalitikus tapasztalat feltárja a számunkra. In: *Thalassa* (4) 1993, 2.

Marosán György 2019: Mi vár ránk a szingularitáson túl? *Népszava,* 2019. 12. 15.

Mori, Masahiro 2012: The uncanny valley. In: *IEEE Robotics and Automation* 19. (2012), 2.

Németh Szabolcs 2021: A közösségi oldalak szolgáltatóinak jogi felelőssége. PhD értekezés (műhelyvitára benyújtott változat). Károli Gáspár Református Egyetem. https://ajk.kre.hu/images/doc2021/doktori/Nemeth_Szabolcs_PhD_dolgozat_muhelyvitara_FINAL.pdf

Péterfalvi Attila – Révész Balázs – Buzás Péter (szerk.) 2018: *Magyarázat a GDPR-ról.* Wolters Kluwer, Budapest.

Pethő Marcell 2021: *Aranyérem a szabályozásban?* Blogbejegyzés. https://www.ludovika.hu/blogok/messzelato/2021/12/08/aranyerem-szabalyozasban/

Pokol Béla 2018: *A mesterséges intelligencia társadalma.* Budapest, Kairosz.

Russell, Stuart J. – Norvig, Peter 2000: *Mesterséges Intelligencia – Modern megközelítésben.* Budapest, Panem.

Szathmáry Zoltán – Miskolczi Barna 2018: *Büntetőjogi kérdések az információk korában (mesterséges intelligencia, big data, profilozás).* Budapest, HVG Orac.

Szepesvári Csaba 2005: *Gépi tanulás – rövid bevezetés.* Előadás. MTA SZTAKI.. 03. 22. http://old.sztaki.hu/~szcsaba/talks/lecture1.pdf

Veale, Michael – Edwards, Lilian 2018: Clarity, surprises and further questions in the Article 29 Working Party draft guidence on automated decision making and profiling. In: *Computer, Law and Security Review* 34 (2018) 2. 398–404.

Z. Karvalics László 2015: Mesterséges intelligencia – a diskurzusok újratervezésének kora. In: *Információs Társadalom* 15., 7–41. http://epa.oszk.hu/01900/01963/00050/pdf/EPA01963_informacios_tarsadalom_2015_4_007-041.pdf

Ződi Zsolt 2021: *A mesterséges intelligencia jogi fogalma.* Blogbejegyzés. https://www.ludovika.hu/blogok/itkiblog/2021/06/18/a-mesterseges-intelligencia-jogi-fogalma/