Az Európai Unió mesterséges intelligencia szabályozása

Szerző: Nagy Demeter Viktor, ABT junior manager

Hosszú várakozás után, április 21-én jelent meg az Európai Unió mesterséges intelligencia rendeletének tervezete, amely az első komolyan vehető, globális hatású próbálkozás az MI alapú rendszerek szabályozására.

Az Európai Uniós jogalkotás kiemelt jelentőséggel bír, mivel a régió hatalmas és egységes piaca igazodásra, a folyamataik valamint kontrollrendszereik felülvizsgálatára készteti a multinacionális vállalatokat, és olyan helyekre is elviszi a jogszabályok hatásait, ahová az Unió jogköre egyébként nem érne el. Ez történt az Európai Unió adatvédelmi rendeletével (GDPR) is, amelynek globális hatásai kimutathatóak, mind a vállalatok viselkedésében, mind más államok jogalkotásában. A világ hasonló hatást vár az MI rendelettől is, ezért élénk figyelem követi az Unió tevékenységét e területen is, bár a rendelet elfogadása és hatálybalépése egyelőre messzinek tűnik. A szöveg végleges változatának európai parlamenti szavazását leghamarabb 2022-re várják, ezt követően – hasonlóan a GDPR-hoz - következik egy kétéves felkészülési idő a hatálybalépést megelőzően. Hatályos MI rendelettel tehát legkorábban 2024 körül érdemes számolni, még úgy is, hogy a szlovén elnökség zászlajára tűzte a digitális átállással kapcsolatos jogszabályok elfogadásának felgyorsítását az ősz folyamán.

Bejegyzésünkben bemutatjuk az Európai Unió MI rendeletének első, nyilvános szövegváltozatát. A terjedelmi korlátok miatt a főbb pontokra fókuszálunk, akit egy szinttel részletesebben is érdekel a jogszabály, annak javasoljuk a Dataprivacy.hu blogon, Bereczki Tamás cikkét vagy a Brookings Institution angol nyelvű írását. Ha kíváncsi mire jutott az Európai Unió, de nem szeretne elmélyedni a témában, a cikk főbb megállapításai a következők:

- **Az EU új mesterséges intelligencia rendelete várhatóan 2024-ben vagy 2025-ben válik hatályossá**, és – hasonlóan a GDPR-hoz –költségeket ró majd azon vállalkozásokra, amelyek az EU-n belül mesterséges intelligencia rendszereket hoznak forgalomba, forgalmaznak vagy használnak fel, különösen, ha azok magas kockázatú MI rendszerek. A Bizottság hatásvizsgálatában úgy becsüli, hogy a **megfelelés költségei a magas kockázatú rendszerek esetében a teljes ár 4-5%-át tehetik majd ki**.
- **Bizonyos MI gyakorlatok (pl. social scoring) teljes tiltása mellett, az EU szigorú követelményeket támaszt majd a magas kockázatú MI rendszerek felé**, vagyis az olyan rendszerek irányába, amelyek alapvető társadalmi működéshez (például: bűnüldözés, oktatás, bevándorlás, igazságszolgáltatás stb.) kapcsolódnak. Ezen túlmenően a tervezet „csak" transzparenciára vonatkozó követelményeket ír elő, azokat is csak az emberi interakcióra szánt rendszerek (pl. chatbotok) esetén.
- Az Európai Unió egy sor olyan intézkedést is javasol, amelyek az európai innovációt hivatottak serkenteni: létrejöhetnek MI tesztkörnyezetek a gépi tanuló algoritmusok egyszerűbb feltanítására, könnyítéseket kaphatnak az MI startupok és létrejöhet a dedikált Európai Mesterséges Intelligencia Hatóság is.

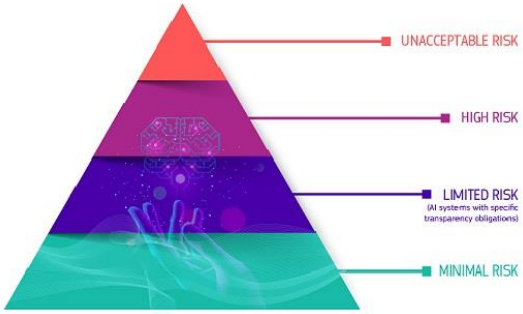
Kifut-e a kezünkből a homok?

A felemelkedőben lévő technológiák meghatározása és szabályozása olyan, mint finomszemű homokot vagy vizet tartani a kezünkben: csak abban lehetünk biztosak, hogy minél tovább várunk, annál kevesebbet, utóbb a semmit markoljuk majd. A mesterséges intelligencia kapcsán, a jogalkotónak nem kizárólag az elmúlt évtizedek eredményeit kellene szabályoznia, hanem keretet is kellene adnia a következő évtized vívmányainak is. Tíz évvel ezelőtt a legtöbb olyan gyakorlat, amelynek használata ma már mindennapos (gépi tanulás, nyelvi feldolgozás), gyakorlatilag csak az innováló vállalatok kutatási-fejlesztési nóvumai között szerepelt. Az analógia alapján nem kétséges, hogy tíz év múlva elterjedt gyakorlatnak számít majd egy sor, ma még csak kipróbálási fázisban lévő technika: a tömeges, MI alapú tartalomgyártás (a la GPT-3, elmozdulás a multimodális modellek felé), élethű chatbotok vagy a közel teljes önvezetés az autózásban.

Az Európai Unió, hogy ezt a csapdát megpróbálja elkerülni, a mesterséges intelligencia pontos meghatározása helyett, - jobbára helyesen - mesterséges intelligencia gyakorlatokat (gépitanuló algoritmusokat, logikai és tudásalapú, valamint statisztikai alapú megközelítéseket) von hatókörbe, konkrét szoftvertípusok helyett. A tervezet definícióját több kritika is érte, elsősorban amiatt, mert a Bizottság az EU saját szakértői csapatának (High-Level Expert Group on Artificial Intelligence – AI HLEG) megközelítése helyett egy teljesen újat írt, amely inkább a közbeszédhez igazodik és a szakma szerint kevésbé alkalmas minden szabályozandó eset körül határolására.

Kockázat-alapú megközelítés

A tervezet az adatvédelmi rendelethez (GDPR) hasonlóan, kockázatalapú megközelítést alkalmaz: az adott MI gyakorlat személyekre, és jogaikra vonatkozó kockázatai alapján állapít meg követelményeket a következő csoportok szerint:



1. Tiltott MI gyakorlatok
2. Magas kockázatú MI gyakorlatok
3. Korlátozott kockázatú MI gyakorlatok
4. Kockázatot nem hordozó MI gyakorlatok

Az Unió kategóriái bár nagyjából fedik a közbeszédben kialakult kockázati szinteket, ugyanakkor egyszerre igyekeznek kezelni a már jelenlévő és még példa nélkül álló eseteket, ami a jövőben még okozhat problémákat.

Tiltott MI gyakorlatnak a rendelet tervezete a következő eseteket tartja:

1. **azon MI gyakorlatok, amelyek finom manipulációs technikák alkalmazása által testi vagy lelki kárt okozhatnak természetes személyeknek**
2. **a hatóságok által vagy hatóságok számára végzett társadalmi megbízhatósági pontozási technikák** (social scoring) azon használatát, amely aránytalanul hátrányosan érint egy személyt vagy csoportokat azok közösségi magatartásához képest, vagy amely más szociális kontextusokba átemelve hátrányos megkülönböztetésre alkalmazza a pontozási technikákat (pl. bűnelkövetői adatbázis létrehozása bűncselekmények előrejelzésére)
3. **a valós idejű biometrikus azonosítórendszerek** (pl. arcfelismerő kamerarendszerek) alkalmazása bűnüldözési célra, meghatározott kivételekkel (eltűnt személyek, valamint valamennyi súlyosabb bűncselekmény gyanúsítottjainak felkutatása, másként el nem hárítható veszély elhárítása), amelyek megkérdőjelezik a tiltás valódi erejét.

Magas kockázatú MI gyakorlatok

A tervezet szövegének tetemes része foglalkozik a magas kockázatú mesterséges intelligencia gyakorlatokra vonatkozó követelményekkel. A magas kockázatú alkalmazásokat a rendelet nem tiltja, de szigorú szabályokat ír elő az alkalmazásukhoz és egy megfelelőségértékelési folyamatnak is aláveti őket, még az Uniós piacra való bevezetésüket megelőzően. Ez utóbbi gyakorlatilag egységesen váltotta ki a piaci szereplők nemtetszését, holott konvencionálisabb már nem is lehetne: a CE tanúsítási rendszerben minden évben termékek milliói esnek át hasonlón. A magas kockázatúnak minősített rendszereket a szolgáltató kizárólag akkor hozhatja forgalomba az Unióban, ha a rendszer átesik egy megfelelőségértékelésen (CE tanúsítási rendszer, hasonlóan egy sor más termékhez), amelynek célja a rendelet tulajdonképpeni követelményeinek ellenőrzése, amelyeket aztán a termék teljes életciklusán át biztosítani kell, és jelentős változás esetén az értékelést is újra el kell végezni.

A tervezet szerint magas kockázatú gyakorlatnak számítanak azon alkalmazások, amikor meghatározott jogszabályok alá eső termékben az MI, mint biztonsági komponens jelenik meg (pl. automata ütközéselkerülő rendszer egy gépjárműben), vagy, ha az alábbi területeken kívánják alkalmazni az MI rendszert:

- Nem hatóságok számára végzett valós idejű biometrikus azonosítás
- Kritikus infrastruktúra irányítása és működtetése
- Foglalkoztatás, a munkavállalók irányítása és az önfoglalkoztatáshoz való hozzáférés
- Alapvető magán- és közszolgáltatásokhoz való hozzáférés
- Bűnüldözés
- Migráció, menekültügy és határellenőrzés
- Igazságszolgáltatás és demokratikus folyamatok ellenőrzése.

A szöveg a magas kockázatú rendszerekre részletes követelményeket állapít meg (a részletezéstől a formai keretek miatt ezúttal eltekintünk), például:

- Bejelentési kötelezettség a központi, Uniós adatbázisba
- Bevezetés utáni nyomon követés és teljesítmény visszamérés
- Súlyos incidensek bejelentése a piacfelügyeleti hatóságnak
- Kockázatmenedzsment rendszer üzemeltetése
- Adatkormányzati (data governance) rendszer üzemeltetése
- Műszaki dokumentáció
- Részletes naplózás
- Átláthatóság és a rendszert felhasználó személy vagy szervezet tájékoztatása
- Végfelhasználó tájékoztatása az MI használat tényéről
- Emberi felügyelet és vészkapcsoló
- Pontosság, robusztusság és kiberbiztonság
- Adatokhoz és dokumentumokhoz való hozzáférés biztosítása a piacfelügyeleti hatóságok számára

Az Európai Bizottság saját hatásvizsgálata úgy számol, hogy a megfelelés költségei az összes MI alkalmazás 5-15%-át jelentő, magas kockázatú rendszerek esetén, a fejlesztési költségek 4-5%-át tehetik majd ki. Fontos látni, hogy a szigorúnak tűnő követelmények egy jelentős része olyan kívánalmat takar, amelyet a fejlesztő vállalatok jelenleg is teljesítenek, mivel:

- iparági jó gyakorlatnak számít (pl. data governance, naplózás)
- gazdasági érdek fűződik hozzá (pl. műszaki dokumentáció, nyomonkövetés)
- jelenleg is jogszabályi követelmény, például a GDPR-ból fakadóan (pl. átláthatóság az automatikus döntéshozatal során).

A CE megfelelőségi rendszer alkalmazása, s még inkább a várható költségek gondolata szinte teljes egységfrontot hozott létre az MI területen érdekelt szereplők között, amely egy sor racionális kritika és – sajnos - csúsztatásokkal terhelt publikáció képében öltött testet.

Korlátozott kockázatú MI gyakorlatok

A tervezet valamennyi, természetes személyekkel való interakcióra szánt MI rendszer esetében előír átláthatósági kritériumokat, melyek szerint elvárás, hogy:

- a természetes személyeket tájékoztassák a használat tényéről (hacsak ez a kontextusból nem nyilvánvaló), ha pedig biometrikus vagy érzelem felismerő rendszerről van szó, akkor a rendszer működéséről is
- a természetes személyeket tájékoztassák az „engedélyezett” deepfake (megtévesztő pontosságú mesterségesen generált kép, hang vagy videótartalom) használata esetén a technika használatáról

A korlátozott kockázatú gyakorlatokra vonatkozó követelmények jó előre szabályozni kívánják az egyre életszerűbbé váló interakciókra képes MI rendszerek kockázatait. Bár a közösségi média platformok eddig nagyon eredményesen vették fel a harcot a deep fakek ellen, s az ipar is csak kacsingat az alkalmazási lehetőségek felé, a technológia egy sor olyan etikai kérdést vet fel, s kockázatot hordoz, amelyre látszólag nincs válasza az emberiségnek. Az elmúlt hónapokban komoly kritikát kapott a néhány éve elhunyt Anthony Bourdainről szóló film, amelynek készítői deep fake technológia segítségével teremtették újra a séf hangját narrációs betétekhez.

Az európai MI innováció útja: MI tesztkörnyezet, Európai Mesterséges Intelligencia Testület

Az Unió mesterséges intelligencia csomagjának a szabályozáson túl, másik fontos célja, hogy az MI versenyben az Egyesült Államok és Kína mögött lemaradni látszó kontinenst felzárkózási pályájára állítsa. A pénzügyi eszközök biztosítása mellett a javaslatcsomag egyéb lépéseket is tartalmaz, így az Európai Mesterséges Intelligencia Testület (European Artificial Intelligence Board) létrehozását, könnyítéseket a kis- és indulóvállalkozások (startupok) számára, a gépi tanuláshoz szükséges személyes adatkezelési környezet megteremtését, valamint az MI szabályozói tesztkörnyezetek létrehozását. Utóbbi két lépés önmagában is jelentős előny lehet az MI gyakorlatokat fejlesztő vállatok számára.

A mesterséges intelligencia alapú rendszerek talán legfontosabb technológiai komponense napjainkra (és a közeljövőre nézve is) a gépitanuló algoritmusok széleskörű alkalmazása. A gépi tanulás során az algoritmusok valós korábbi példákból tanulva pontosítják saját modelljeiket és így pontosságukat is. A gépi tanuláshoz hatalmas adattömegekre van szükség, amelyek hozzáférhetősége (vagy inkább hozzáférhetetlensége) az érintett rendszerek fejlődésének jelentős gátja. Ezt az akadályt hivatott a jogalkotó elhárítani azzal, hogy

- lehetőséget biztosít a tagállami és európai illetékes hatóságoknak az MI szabályozói tesztkörnyezetek létrehozására
- lehetőséget biztosít a más célból gyűjtött személyes adatok tesztkörnyezetekben történő felhasználására, kontrollált körülmények között, amennyiben az jelentős közérdeket szolgál (pl. közbiztonság, közegészségügy, bűnmegelőzés).

A fenti könnyítések azt jelentik, hogy egy egészségügyi startup lehetőséget kaphat arra, hogy innovatív MI rendszerét például a magyar egészségügy bizonyos adatsorai segítségével tanítsa és ezzel hozzájáruljon súlyos megbetegedések előrejelzéséhez. Bár a szabályozói tesztkörnyezetek a fintech megoldások terén már bevett módszernek számítanak, széles körű alkalmazásukra az MI technológiák segítésére most kerül sor először (kisebb léptékben, a norvég adatvédelmi hatóságnak már van működő projektje és az angol ICO is használ adatvédelmi tesztkörnyezetet), ami egyedülálló előnyhöz juttathatja a résztvevő vállalatokat és ezzel együtt az Európai Uniót is.

Milyen hatással lehet ez ránk?

Bár hatályos és alkalmazandó mesterséges intelligencia rendeletre legkorábban 2024 körül érdemes számítani, a felelős szervezeti magatartás előrelátást és proaktivitást diktál. Amennyiben a rendelet végleges követelményei élesednek az MI alapú rendszereket fejlesztő vagy használó szervezetek hirtelen egy sor feladattal találják majd szemben magukat. Aki emlékszik az utolsó nagy, digitális technológiát érintő európai reguláció időszakára, a GDPR bevezetésére az jó eséllyel fel tudja mérni a várható feladatok nagyságát. Még inkább igaz ez, ha számba vesszük, hogy bár egységes szabályozás még nincs, MI-re vonatkozó szabályok már jelenleg is vannak, jól lehet a jogalkalmazói akart hiányzik a kikényszerítésükre. Adatvédelmi oldalról a GDPR szigorú szabályokat ír elő az automatikus döntéshozatal bizonyos eseteire, de a második világháború utáni nyugati jogalkotás nagy vívmányai, a termékbiztonsági, fogyasztóvédelmi és diszkriminációt tiltó rendelkezések is kötelmek sorát rójják az MI alkalmazásokra, például a diszkrimináció tiltásával, amelynek jelenléte korunk nagy MI modelljeinek egyik legnagyobb hibája. Amennyiben a jogalkalmazói visszafogottság oka megszűnik, a szervezetek egyszerre szembesülhetnek az új, egységes szabályozás és a korábbi, fragmentált

kötelmek kihívásaival is. Bár a mesterséges intelligencia adaptáció még kezdeti fázisában jár (Magyarországon például az EuroStat adatai szerint a pénzügyi szektoron kívüli, tíz főnél több munkavállalót foglalkoztató vállalatok mindössze három százaléka használ MI alkalmazásokat), a jövőben gyors növekedés várható a területen. Bár a globális szabályozások még váratnak magukra, a piac és a vásárlók előreláthatóan sokkal hamarabb kikényszerítik majd a specifikus kontrollrendszerek és etikus MI gyakorlatok alkalmazását. Amennyiben vállalkozása fontolgatja mesterséges intelligencia alkalmazás használatát, úgy – konzultatív jelleggel – érdemes figyelembe vennie az európai szabályozási környezet várható alakulását is, valamint algoritmus etikai megfontolásokat is.